1 前言

§1.1 动力

想起来是那样遥远——从Lebesgue发表他的建立我们现在所称的Lebesgue测度与积分的几页纸的划时代的论文算起, 整整一百年过去了. 一百年春风秋雨, 一百年春华秋实, Lebesgue的工作经受住了时间和实践的检验——这个当年虽然茁壮但却幼小的禾苗长成了枝繁叶茂的大树——它的枝叶荫及了现代数学中的许多领域. 回望来路, 我们可以看到许多大大小小的数学家在树下辛勤劳作的身影; 瞻望未来, 它将长久地为后来者提供庇荫.

人们为什么需要它? 有无穷多条理由. 我们只要看一下概率论的情况就可以了.

初等概率论中有许多没有严格证明的结论。例如, 我们都学过下面这个结论：设ξ为一连续型随机变量，分布密度为f, 则对于任意Borel可测函数F, $F(\xi)$也为随机变量且

$$E[F(\xi)] = \int F(x)f(x)dx.$$

但一般教科书并没有给出这个结论的证明, 因为只凭初等概率的知识是无法证明的. 再如，我们学过独立随机变量序列的种种性质，但却"忽略"了一个基本问题：这种序列是否存在？如果它们根本就不存在的话，我们所学的一切岂不是全是空中楼阁？当然这并不是粗心所造成的忽略，而是在初等概率论中，根本就无法回答这些问题。

而有了测度论, 就可以回答这些问题. 当然, 历史的发展证明, 测度论对概率论的作用远不止于此. 实际上，整个现代概率论的基础，全是建立在测度论基础之上的。这一工作是20世纪30年代由上世纪最伟大的科学家之一——前苏联数学家Kolmogorov完成的.

高等概率论这门课的主要内容，就是测度论.

§1.2 计划

这本书主要是写给概率统计方面的研究生的入门教材和研究人员的参考书. 它无意成为一部内容齐全的辞书，内容的选取无疑地是受到作者本人的经验的影响的.

2 集类与可测性

§2.1 基本术语

我们已经知道了集合论的基本概念，也知道了下列运算：

1、并：

$$E_1 \cup E_2, \qquad\qquad \cup_{i\in I} E_i$$

2、交：

$$E_1 \cap E_2, \qquad\qquad \cap_{i\in I} E_i$$

以后我们还要经常用到以下概念。
3、特征函数

$$1_E(x) = 1 \text{ if } x \in E, 0 \text{ if } x \notin E.$$

集合与其特征函数相互唯一确定。
现在设X为一集合，$E_n \subset X$, $n = 1, 2, \cdots$.
4、上极限

$$\begin{aligned}\limsup_n E_n : &= \{x \in X : \exists\text{无穷多个}n\text{使}x \in E_n\} \\ &= x \in X : \forall n, \exists k \geq n \text{ 使}x \in E_k \\ &= \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k.\end{aligned}$$

5、下极限

$$\begin{aligned}\liminf_n E_n : &= \{x \in X : \text{只有有穷多个}n\text{使}x \notin E_n\} \\ &= x \in X : \exists n, \forall k \geq n x \in E_k \\ &= \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} E_k.\end{aligned}$$

6、极限
若$\limsup_n E_n = \liminf_n E_n$，则称极限存在，且记为：$\lim_n E_n$.
7、单调列
若$\forall n$, $E_n \subset E_{n+1}$, 则称为单调上升列,记为$E_n \uparrow$. 此时$\lim_n E_n = \cup_n E_n$.
若$\forall n$, $E_n \supset E_{n+1}$, 则称为单调下降列,记为$E_n \downarrow$. 此时$\lim_n E_n = \cap_n E_n$.
8、余集

$$E^c := \{x : x \notin E\}, \qquad\qquad E \subset F \Leftrightarrow E^c \supset F^c.$$

9、De Morgan原理

$$(\cup_{i\in I} E_i)^c = \cap_{i=1} E_i^c,\ (\cap_{i\in I} E_i)^c = \cup_{i=1} E_i^c.$$

$$F - \cup_{i\in I} E_i = \cap_{i=1}(F - E_i),\ (F - \cap_{i\in I} E_i) = \cup_{i=1}(F - E_i).$$

10、差

$$E - F := E \cap F^c.$$

11、对称差

$$E \triangle F := (E - F) \cup (F - E).$$

12、运算法则

$$A \cap B = B \cap A, \quad A \cup B = B \cup A$$

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

$$(A \cap B) \cap C = A \cap (B \cap C)$$

$$(A \cup B) \cup C = A \cup (B \cup C)$$

$$\cup_{i \in I} A_i - B = \cup_{i \in I}(A_i - B)$$

§2.2 常用集类

我们考虑问题时，总会局限在一定的范围内。随着问题的变化，这个范围也会变化。但当一个问题确定下来后，这个问题一般也会随之确定下来。这样的一个范围，我们叫它空间，或有时为了强调其大而称为全空间。例如，我们考虑一维问题时，这个空间是$\mathbb{R}^1$；考虑n维问题时，它是$\mathbb{R}^n$，等等。以后如不特别申明，我们都假定已预设了这样一个全空间：一切元素、子集等都是这个空间的。我们暂时用X表示这个空间.

由X的某些子集所构成的集合称为集类。我们以后要研究的集类全都满足某种运算性质. 第一种集类是: 环.

定义2.2.1. 设$\mathscr{R}$为非空集类。若

$$E, F \in \mathscr{R} \Rightarrow E \cup F, \qquad E - F \in \mathscr{R},$$

则称$\mathscr{R}$为环。

注：$\forall$环$\mathscr{R}, \emptyset \in \mathscr{R}$. $(\because \emptyset = E - E \in \mathscr{R})$.
什么是环?

例2.2.2. : 设X是一集合，则X的所有有限子集构成的集类是环。

例2.2.3. : $\mathscr{R} := \{\cup_{i=1}^n [a_i, b_i), n = 1, 2, \cdots, -\infty < a_i \leq b_i < +\infty\}$是环。

证明：易见$\mathscr{R}$对有限交与并均封闭。又因为

$$[a,b)-[c,d)\in\mathscr{R},$$

所以由De Morgan法则

$$\begin{aligned}
& \cup_{i=1}^n[a_i,b_i)-\cup_{j=1}^m[c_j,d_j) \\
= \ & \cup_{i=1}^n([a_i,b_i)-\cup_{j=1}^m[c_j,d_j)) \\
= \ & \cup_{i=1}^n\cap_{j=1}^m([a_i,b_i)-[c_j,d_j)) \\
\in \ & \mathscr{R}.
\end{aligned}$$

Q.E.D.

什么不是环？

例2.2.4. : 设X是一无穷集合，n为一整数，则X的所有由n个元素组成的有限子集构成的集类不是环。

例2.2.5. $\mathscr{R}:=\{[a,b),-\infty<a<b<+\infty\}$不是环。

命题2.2.6. $\mathscr{R}$ 为环$\Leftrightarrow$ $\mathscr{R}$关于并及真差封闭。

证明：$\Rightarrow$显然。$\Leftarrow$:设$E,F\in\mathscr{R}$, 则$E\cup F\in\mathscr{R}$, 因而$E-F=E\cup F-F\in\mathscr{R}$.

命题2.2.7. 设$\mathscr{R}$为环，则

1. $E,F\in\mathscr{R}\Rightarrow E\triangle F,E\cap F\in\mathscr{R}$.

2. $E_i\in\mathscr{R}$, $i=1,2,3,\cdots\Longrightarrow\cup_{i=1}^nE_i\in\mathscr{R},\qquad\cap_{i=1}^nE_i\in\mathscr{R}$.

证明：1.

$$E\triangle F=(E-F)\cup(F-E),$$

$$E\cap F=(E\cup F)-(E\triangle F).$$

2. 归纳法。

定义2.2.8. 代数：环$\mathscr{R}$若含全空间，则称为代数。

上面的例2.2.3只是环而不是代数。反之，下面的集类是代数：

$$\begin{aligned}
\mathscr{A}:=\{\quad & \cup_{i=1}^nE_i,n=1,2,\cdots,\ \text{其中}E_i\ \text{或者为}[a,b)-\infty<a<b\le+\infty, \\
& \text{或者为}(-\infty,b),b\le+\infty\}.
\end{aligned}\tag{2.2.1}$$

要验证一个集类为代数，当然可以直接用定义，先证明它为环，再验证它包含全空间。但通过下面的结果去验证会更快捷一些。

命题2.2.9. 一非空集类$\mathscr{R}$为代数的充要条件是下面两条之一（因而同时）成立:

(1) $\mathscr{R}$对并及余封闭;

(2) $\mathscr{R}$对交及余封闭。

证明: 两个条件的必要性显然, 往证充分性. 设全空间为X.

1. 设$\mathscr{R}$对并及余封闭, 首先注意任取$F \in \mathcal{R}$, 有$F^c \in \mathcal{R}$, 从而

$$\Omega = F \cup F^c \in \mathscr{R}.$$

其次,设$F, E \in \mathscr{R}$, 则

$$E - F = E \cap F^c = (E^c \cup F)^c \in \mathscr{R}$$

所以$\mathscr{R}$是代数. 同理可证(2).

Q.E.D.

§2.3 生成环与代数

我们仍然假定全空间为X, 一切集类均是这个空间的子集所构成的集类.

定义2.3.1. 设$\mathscr{E}$为一集类，$\mathscr{R}_0$为环, $\mathscr{A}_0$为代数.

1. 若$\mathscr{E} \subset \mathscr{R}_0$且对任意包含$\mathscr{E}$的环$\mathscr{R}$ 有$\mathscr{R}_0 \subset \mathscr{R}$, 则称$\mathscr{R}_0$为$\mathscr{E}$所生成的环，记为$\mathscr{R}_0 = \mathscr{R}(\mathscr{E})$.

2. 同样地, 若$\mathscr{E} \subset \mathscr{A}_0$且对任意包含$\mathscr{E}$的代数$\mathscr{A}$ 有$\mathscr{A}_0 \subset \mathscr{A}$, 则称$\mathscr{A}_0$为$\mathscr{E}$所生成的代数，记为$\mathscr{A}_0 = \mathscr{A}(\mathscr{E})$.

定理2.3.2. $\forall$集类$\mathscr{E}$, $\mathscr{R}(\mathscr{E})$与$\mathscr{A}(\mathscr{E})$均存在。

证明: 以环为例. 以$\mathscr{P}(X)$表示X的所有子集构成的集类，则$\mathscr{P}(X)$为环且$\mathscr{P}(X) \supset \mathscr{E}$.

于是，令

$$\mathscr{Q} := \{\mathscr{R} : \mathscr{R} \text{ 为环且} \mathscr{R} \supset \mathscr{E}\},$$

则$\mathscr{Q}$非空。

令

$$\mathscr{R}_0 := \cap_{\mathscr{R} \in \mathscr{Q}} \mathscr{R} = \{E : \forall \mathscr{R} \in \mathscr{Q}, E \in \mathscr{R}\}.$$

则$\mathscr{R}_0$为环且$\forall \mathscr{R} \in \mathscr{Q}$, $\mathscr{R}_0 \subset \mathscr{R}$. $\therefore \mathscr{R}_0 = \mathscr{R}(\mathscr{E})$

推论2.3.3. 设$\mathscr{E}$为集类，则$\forall A \in \mathcal{R}(\mathscr{E})$, $\exists n$, $B_i \in \mathscr{E}$, $i = 1, 2, \cdots, n$, 使$A \subset \bigcup_{i=1}^n B_i$.

注意对生成的代数, 该定理不成立.

证明：令

$$\mathscr{T} := \{A : \exists n, \exists B_i, i = 1, 2, \cdots, n, \text{ 使} A \subset \bigcup_{i=1}^{n} B_i\}.$$

则$\mathscr{T}$为环且$\supset \mathscr{E}$. $\therefore \mathscr{R}(\mathscr{E}) \subset \mathscr{T}$.

例2.3.4.

$$X = \mathbb{R}, \quad \mathscr{E} = \{[a, b), -\infty < a \leq b < +\infty\}$$

$$\mathscr{R}(\mathscr{E}) = \{\cup_{i=1}^{n}[a_i, b_i), n = 1, 2, 3, \cdots, -\infty < a_i \leq b_i < +\infty\}$$

$$\begin{aligned}\mathscr{A}(\mathscr{E}) = \{ \quad & \cup_{i=1}^{n} E_i, n = 1, 2, \cdots, \text{ 其中} E_i \text{ 或者为}[a, b) - \infty < a \leq b \leq +\infty, \\ & \text{或者为}(-\infty, b), b \leq +\infty\}.\end{aligned}$$

例2.3.5. 设X为任一空间, $\mathscr{E} = \{\{x\}, x \in X\}$, 则

$$\mathscr{R}(\mathscr{E}) = \{A : A \text{ 为} X \text{的有限子集}\}$$

$$\mathscr{A}(\mathscr{E}) = \{A : A \text{ 或} A^c \text{ 为} X \text{的有限子集}\}$$

注：

$$\mathscr{R}(\mathscr{E}) = \mathscr{R}(\mathscr{R}(\mathscr{E})).$$

$$\mathscr{A}(\mathscr{E}) = \mathscr{A}(\mathscr{A}(\mathscr{E})) = \mathscr{A}(\mathscr{R}(\mathcal{E})).$$

§2.4 σ-代数与可测空间

对极限运算封闭的代数称为σ代数.

定义2.4.1. 非空集类$\mathscr{F}$若满足

*1.*对余封闭：

$$E \in \mathscr{F} \Longrightarrow E^c \in \mathscr{F};$$

*2.*对可列并封闭：

$$E_i \in \mathscr{F}, i = 1, 2, \cdots \Longrightarrow \bigcup_{i=1}^{\infty} E_i \in \mathscr{F},$$

则称为(X上的子集)σ代数.

σ代数当然是代数. 而如果已经知道一个集类是代数, 则验证它是σ代数的手续要简单一些.

命题2.4.2. 代数$\mathscr{F}$如果对不交的可列并封闭, 则成为σ代数.

证明: 设$E_n \in \mathscr{F} \quad n \geq 1$. 补充定义$E_0 = \emptyset$. 由于$\mathscr{F}$是代数, 故$F_n := E_n - \cup_{k=0}^{n-1} E_k \in \mathscr{F}$. 但$F_n, n \geq 1$不交, 故

$$\cup_{n=1}^{\infty} E_n = \cup_{n=1}^{\infty} F_n \in \mathscr{F}.$$

完毕.

命题2.4.3. 设$\mathscr{F}$为σ代数, $E_i \in \mathscr{F}$, 则

$$\cap E_i, \quad \limsup E_i, \quad \liminf E_i \in \mathscr{F}.$$

证明:

$$\because E_i \in \mathscr{F}, \therefore E_i^c \in \mathscr{F}, \therefore \cup_{i=1}^{\infty} E_i^c \in \mathscr{F}, \therefore (\cup_{i=1}^{\infty} E_i^c)^c \in \mathscr{F},$$

即

$$\cap_{i=1}^{\infty} E_i \in \mathscr{F}.$$

进而

$$\limsup E_i = \cap_{i=1}^{\infty} \cup_{k=i}^{\infty} E_k \in \mathscr{F},$$

$$\liminf E_i = \cup_{i=1}^{\infty} \cap_{k=i}^{\infty} E_k \in \mathscr{F}.$$

定理2.4.4. 设$\mathscr{E}$为一非空集类, 则存在$\sigma-$代数$\mathscr{F}_0$使

1. $\mathscr{E} \subset \mathscr{F}_0$.

2. 若$\mathscr{F}$为$\sigma-$代数且$\mathscr{F} \supset \mathscr{E}$, 则$\mathscr{F} \supset \mathscr{F}_0$.

$\mathscr{F}_0$称为由$\mathscr{E}$产生的$\sigma-$代数, 记为$\sigma(\mathscr{E})$.

证明与定理2.3.2的证明类似,留作练习。

注:对任意集类$\mathscr{E}$,

$$\sigma(\mathscr{E}) = \sigma(\sigma(\mathscr{E}))$$

$$\sigma(\mathscr{E}) = \sigma(\mathscr{A}(\mathscr{E})).$$

例2.4.5.

$$X = \mathbb{R}, \quad \mathscr{E} = \{[a, b), -\infty < a \leq b < +\infty\}$$

我们知道

$$\mathscr{A}(\mathscr{E}) = \{ \quad \cup_{i=1}^{n} E_i, n = 1, 2, \cdots, \text{ 其中} E_i \text{ 或者为}[a, b) - \infty < a \leq b \leq +\infty, \text{或者为}(-\infty, b), b \leq +\infty\}.$$

但$\mathscr{A}(\mathscr{E})$不是σ代数. 例如,若

$$A_n = [\frac{1}{n}, 1)$$

则

$$\cup_{n=1}^{\infty} A_n = (0, 1) \notin \mathscr{A}(\mathscr{E}).$$

所以

$$\mathscr{A}(\mathscr{E}) \subset \sigma(\mathscr{E}),$$

且

$$\mathscr{A}(\mathscr{E}) \neq \sigma(\mathscr{E}).$$

那么$\sigma(\mathscr{E})$包含了什么元素呢?

首先,对任意开区间(a, b),

$$(a, b) = \cup_{n=1}^{\infty} [a + \frac{1}{n}, b) \in \sigma(\mathscr{E}).$$

又由于直线上的任意开集均可表示为可数个开区间的并,所以$\sigma(\mathscr{E})$包含了所有开集. 进而,由于$\sigma(\mathscr{E})$对余封闭,所以它又包含了所有闭集. 特别地,它包含了任一单点集$\{x\}, x \in \mathbb{R}$. 这样,它还包含了

$$(a, b] = (a, b) \cup \{b\} \quad \forall a, b \in \mathbb{R}.$$

$$[a, b] = (a, b) \cup \{b\} \cup \{a\} \quad \forall a, b \in \mathbb{R}.$$

又由于

$$(-\infty, a) = \cup_{n=1}^{\infty} (-n, a)$$

它包含了$(-\infty, a)$. 同理,它还包含了$(-\infty, a]$, $(a, +\infty)$, $[a, +\infty)$, $\forall a \in \mathbb{R}$. 自然,它也包含了全空间$\mathbb{R} = (-\infty, +\infty)$.

$\sigma(\mathscr{E})$称为$\mathbb{R}$上的*Borel*代数[1], 记为$\mathscr{B}(\mathbb{R})$或$\mathscr{B}_1$(下标"1"代表"一维"), 其元素称为*Borel*集.

习题: 令

$$\mathscr{E}_1 := \{(-\infty, a), a \in \mathbb{R}\}$$

$$\mathscr{E}_2 := \{(-\infty, a], a \in \mathbb{R}\}$$

$$\mathscr{E}_3 := \{(a, b), a, b \in \mathbb{R}\}$$

$$\mathscr{E}_4 := \{(a, b], a \in \mathbb{R}\}$$

$$\mathscr{E}_5 := \{[a, b], a, b \in \mathbb{R}\}$$

[1]Émile Borel, 1871-1956, 法国数学家

$$\mathscr{E}_6 := \{(a, +\infty], a \in \mathbb{R}\}$$

$$\mathscr{E}_7 := \{[a, +\infty), a \in \mathbb{R}\}$$

$$\mathscr{E}_8 := \mathbb{R}\text{中的所有开集}$$

$$\mathscr{E}_9 := \mathbb{R}\text{中的所有闭集}$$

证明: $\forall i = 1, \cdots, 9,$

$$\sigma(\mathscr{E}_i) = \mathscr{B}_1.$$

习题: 将上述构造Borel代数的方法推广到$\mathbb{R}^n$.

习题: 可分度量空间上的Borel代数. 设(X, d)为一可分度量空间, 由X的所有开集产生的σ代数称为$Borel$代数, 记为$\mathscr{B}(X)$. 证明:

1. $\mathscr{B}(X) = \sigma(\text{所有闭集})$;
2. $\mathscr{B}(X) = \sigma(\text{所有开球})$;
3. $\mathscr{B}(X) = \sigma(\text{所有闭球})$;
4. 存在可数个开球生成$\mathscr{B}(X)$;
5. 存在可数个闭球生成$\mathscr{B}(X)$.

记号：设$\mathscr{E}$是集类，A是集合，属不属于$\mathscr{E}$都行。记

$$\mathscr{E} \cap A := \{E \cap A : E \in \mathscr{E}\}.$$

定理2.4.6.

$$\sigma(\mathscr{E}) \cap A = \sigma(\mathscr{E} \cap A).$$

证明：首先，$\sigma(\mathscr{E}) \cap A$为$\sigma-$代数(why?), 又$\mathscr{E} \cap A \subset \sigma(\mathscr{E}) \cap A$,

$$\therefore \sigma(\mathscr{E} \cap A) \subset \sigma(\mathscr{E}) \cap A.$$

反过来，令

$$\mathscr{F} := \{E : E \in \sigma(\mathscr{E}), E \cap A \in \sigma(\mathscr{E} \cap A)\}.$$

则$\mathscr{F}$为$\sigma-$代数且$\mathscr{F} \supset \mathscr{E}$, 因而$\mathscr{F} \supset \sigma(\mathscr{E})$. 故$\mathscr{F} = \sigma(\mathscr{E})$. 于是

$$\sigma(\mathscr{E} \cap A) \supset \sigma(\mathscr{E}) \cap A.$$

Q.E.D.

注意我们证明上述第二步的方法: 为证明具有某种性质的集类(这里是$\mathscr{F}$)包含了另一集类($\mathscr{E}$)产生的σ代数,只要说明前者是σ代数并且包含了后者. 这一方法在测度论中经常使用.

§2.5 可测函数

由于我们以后常常会遇到取值包括$\pm\infty$的函数, 因此将上述定义推广一下会带来许多方便. 即: 将$\{-\infty\}$, $\{+\infty\}$也称为Borel集. 因此我们以后将这两个集合和$\mathscr{B}(\mathbb{R}^1)$共同产生的$\sigma$代数也称为(扩张)Borel代数并以$\mathscr{B}(\overline{\mathbb{R}})$表示之, 即:

$$\mathscr{B}(\overline{\mathbb{R}}) := \sigma(\mathscr{B}(\mathbb{R}), \{-\infty\}, \{+\infty\}).$$

其次我们要引进可测空间的概念.

定义2.5.1. 设X是一空间, $\mathscr{F}$是其子集σ代数. 称二元体$(X, \mathscr{F})$为可测空间.

设X, Y为两空间, $f : X \mapsto Y$为映射, 对$A \subset Y$, 记

$$f^{-1}(A) := \{x \in X : f(x) \in A\},$$

称为f的逆映射. 又若$\mathscr{Y}$为Y上集类,则记

$$f^{-1}(\mathscr{Y}) := \{f^{-1}(A) : A \in \mathscr{Y}\}.$$

注意,容易验证逆映射保持所有的集合运算不变,即

$$f^{-1}(\cap_{i\in I} A_i) = \cap_{i\in I} f^{-1}(A_i),$$

$$f^{-1}(\cup_{i\in I} A_i) = \cup_{i\in I} f^{-1}(A_i),$$

$$f^{-1}(A - B) = f^{-1}(A) - f^{-1}(B),$$

等等. 这是一条十分有用的性质. 特别是我们可以由此推出,对任意集类$\mathscr{A}$, 有

$$f^{-1}(\sigma(\mathscr{A})) = \sigma(f^{-1}(\mathscr{A})).$$

现在可以定义可测函数了.

定义2.5.2. 设$(X, \mathscr{F})$, $(Y, \mathscr{G})$为可测空间, $f : X \mapsto Y$若满足$f^{-1}(\mathscr{G}) \subset \mathscr{F}$, 则称为可测映射, 记为$f \in \mathscr{F}/\mathscr{G}$. 特别地,若$f : X \to \overline{\mathbb{R}^1}$为可测映射,则称为可测函数.

命题2.5.3. 设$(X, \mathscr{F})$, $(Y, \mathscr{G})$为可测空间, $\mathscr{A} \subset \mathscr{G}$且$\sigma(\mathscr{A}) = \mathscr{G}$. $f^{-1} : X \mapsto Y$. 若$f^{-1}(\mathscr{A}) \subset \mathscr{F}$, 则$f \in \mathscr{F}/\mathscr{G}$.

证明: 在所给条件下有

$$f^{-1}(\mathscr{G}) = f^{-1}(\sigma(\mathscr{A})) = \sigma(f^{-1}(\mathscr{A})) \subset \sigma(\mathscr{F}) = \mathscr{F}.$$

推论2.5.4. $f : X \to \overline{\mathbb{R}^1}$为可测函数的充要条件是:

$$\{x \in X : f(x) < a\} \in \mathscr{F}, \qquad \forall a \in \mathbb{R},$$

还可以给出许多类似的充要条件,例如将上式中的严格不等号换为不严格的不等号, 等等.

例2.5.5. *若$(X,\mathscr{F})=(I,I\cap\mathscr{B}(\mathbb{R}))$, 其中$I$为任一区间，则$I$上的所有连续函数均为可测函数.*

这是因为, 若f连续, E是开集, 则$f^{-1}(E)$也是I中的相对开集.

例2.5.6. *设$(X,\mathscr{F})$为可测空间, $A\in\mathscr{F}$, 则$f(x):=1_A(x)$为可测函数.*

因为

$$\{f<a\}=\begin{cases}A^c, & a\le 1\\ X, & a>1\end{cases}\tag{2.5.2}$$

例2.5.7. *若$(X_i,\mathscr{F}_i)$, $i=1,2,3$, 为可测空间, $f\in\mathscr{F}_1/\mathscr{F}_2$, $g\in\mathscr{F}_2/\mathscr{F}_3$, 则$g\circ f\in\mathscr{F}_1/\mathscr{F}_3$.*

证明: $(g\circ f)^{-1}(\mathscr{F}_3)=f^{-1}(g^{-1}(\mathscr{F}_3))\subset f^{-1}(\mathscr{F}_2)\subset cF_1$.

我们知道, 连续函数对一些常见的运算如和、差、积、商（在分母非零时）、复合等都是封闭的. 我们马上会看到, 可测函数也有类似的性质. 不仅如此, 它们对一些关于连续函数不封闭的运算如极限也是封闭的. 从这个意义上讲, 可测函数之于连续函数正如实数之于有理数. 可测函数的重要性由此可见一斑——能想象吗? 我们这个世界如果没有实数! 我们的测度论如果没有下面这个定理!

定理2.5.8. *设$(X,\mathscr{F})$为可测空间, f,g,f_n为可测函数, c是常数. 假定下面出现的所有运算均有意义. 则$f+g$, cf, f^{-1}, fg, $\vee_{n=1}^{\infty}f_n$, $\wedge_{n=1}^{\infty}f_n$, $\limsup_n f_n$, $\liminf_n f_n$均是可测函数.*

这里$\bar{\mathbb{R}}$中的运算作如下规定: 对$x\in\mathbb{R}$,

$$(\pm\infty)+x=x+(\pm\infty)=x-(\mp\infty)=\pm\infty$$

$$(\pm\infty)+(\pm\infty)=(\pm\infty)-(\mp\infty)=\pm\infty$$

$$\frac{x}{\pm\infty}=0$$

$$x\cdot(\pm\infty)=(\pm\infty)\cdot x=\pm\infty(x>0),\quad 0(x=0),\quad \mp\infty(x<0)$$

下面的运算无意义

$$\frac{x}{0},\quad \frac{\infty}{\infty},$$

$$(\pm\infty)-(\pm\infty),\quad (\pm\infty)+(\mp\infty)$$

定理的证明：我们一个一个地证明.

1. 以$\mathbb{Q}$表$\mathscr{B}(\mathbb{R})$中全体有理数的集合，则由于有理数全体可数, 故$\{f+g<a\}=\cup_{r\in\mathbb{Q}}(\{f<r\}\cap\{g<a-r\})\in\mathscr{F}$, $\forall a\in\mathscr{R}$, 即$f+g$可测.

2. 再由于

$$\{cf<a\}=\left\{\begin{array}{l}\{f<c^{-1}a\} \text{ 若} c>0;\\ \{f>c^{-1}a\} \text{ 若} c<0,\end{array}\right.$$

故cf可测.

3. 由以上两点特别得到$f+g$与$f-g$可测. 再注意$x\mapsto x^2$为连续函数, 由例2.5.7知$(f+g)^2$与$(f-g)^2$均可测. 这样,

$$fg=\frac{1}{4}\{(f+g)^2-(f-g)^2\}$$

便可测了.

4. $\{\vee_{n=1}^{\infty}f_n\le a\}=\cap\{f\le a\}$.

5. $\{\wedge_{n=1}^{\infty}f_n\ge a\}=\cup_{n=1}^{\infty}\{f_n\ge a\}$.

6. 先假设$f_n\uparrow$. 由于

$$\lim_n f_n=\vee_{n=1}^{\infty}f_n,$$

故$\lim_n f_n$可测. 同理对单调下降的序列结论也成立.一般情形可由

$$\liminf_n f_n=\lim_n \wedge_{k\ge n}f_k,$$

及

$$\limsup_n f_n=\lim_n \vee_{k\ge n}f_k,$$

得到. 完毕.

最简单的可测函数自然是可测集的特征函数, 即形如1_E的函数，其中$E\in\mathscr{F}$; 稍微复杂一点的是这些函数的线性组合, 它们称为简单函数, 其一般形式是

$$f=\sum_{i=1}^{n}a_i1_{E_i}. \tag{2.5.3}$$

对$\cup_{i=1}^n E_i$做如下分割: 对$A\subset\{1,2,\cdots,n\}$, $A\ne\emptyset$令

$$F_A:=(\cap_{i\in A}E_i)\cap(\cap_{i\notin A}E_i^c)$$

则这样的F_A只有有限个(2^n-1个)且若$A\ne B$, 则$F_AF_B=\emptyset$. 于是

$$f=\sum_A(\sum_{i\in A}a_i)1_{F_A}. \tag{2.5.4}$$

这样我们便说明了简单函数的表达式中, 诸集合是可以选为互不相交的. 于是今后有必要时, 我们就直接做此假定.

有限个简单函数的线性组合、乘积、极大值、极小值都是简单函数.

由上一定理, 简单函数的上、下极限是可测函数. 有用的是, 反过来也是对的.

命题2.5.9. *任一可测函数均是简单函数序列的极限.*

证明: 设f非负可测, 令

$$f_n(x) := \begin{cases} \frac{i-1}{2^n} \text{ 若} \frac{i-1}{2^n} \le f(x) < \frac{i}{2^n} \\ n \text{ 若} f(x) \ge n. \end{cases} \tag{2.5.5}$$

则$f_n \uparrow f$. 一般情况用分解$f = f^+ - f^-$即可. 完毕.

这一定理给我们提供了一个由简单过渡到复杂的途径, 以后我们会看到, 许多构造和结果都是通过这种过渡实现的. 下面就是一个简单而有用的例子.

定理2.5.10. *设X是一空间, $(Y, \mathscr{G})$为可测空间, $F : X \mapsto Y$. 令*

$$\sigma(F) := F^{-1}(\mathscr{G}).$$

则对$(X, \sigma(F))$上的任一可测函数f, 存在$(Y, \mathscr{G})$上的可测函数g使得

$$f = g \circ F.$$

§2.6 单调类定理

一般说来，要验证一个给定的集类是$\sigma-$代数, 到目前为止我们只基本上知道根据定义去验证,而这往往是是相当困难的. 类似地, 一般说来, 要知道一个集类产生的$\sigma-$代数有多大也是相当困难的。但是，对两种比较特殊的集类(代数和π类)，人们还是找到了很好的办法. 事实上, 这些办法决不只是技巧上的小聪明是改进, 而是关乎整个测度论的奠基性的方法: 要是没有它们,我们对很多问题恐怕都是束手无策的.

这就是我们将要介绍的单调类定理. 我们先定义单调类.

定义2.6.1. *非空集类$\mathcal{M}$若满足*

$$\{E_n\} \subset \mathcal{M}, E_n \text{ 单调} \Longrightarrow \lim E_n \in \mathcal{M},$$

则称为单调类.

按定义,σ代数都是单调类,但单调类未必都是σ代数. 不过我们有:

定理2.6.2. *若$\mathscr{F}$既是代数又是单调类, 则它是σ代数.*

事实上, 设$E_i \in \mathscr{F}, i=1,2,\cdots$, 要证明$\cup_{i=1}^{\infty} E_i \in \mathscr{F}$. 为此，我们写它为

$$\cup_{i=1}^{\infty} E_i = \cup_{n=1}^{\infty} \cup_{i=1}^{n} E_i.$$

因为$\mathscr{F}$是代数，故$\cup_{i=1}^{n} E_i \in \mathscr{F}$; 又因为它是单调类，故$\cup_{n=1}^{\infty} \cup_{i=1}^{n} E_i \in \mathscr{F}$.

与环、$\sigma-$代数的情形一样，可以证明

定理2.6.3. *对任意集类$\mathscr{E}$, 存在包含$\mathscr{E}$的最小的单调类，称为$\mathcal{E}$生成的单调类，记为$\mathscr{M}(\mathscr{E})$.*

由于σ代数都是单调类,因此一般说来$\mathscr{M}(\mathscr{E}) \subset \sigma(\mathscr{E})$. 下面就是单调类定理.

定理2.6.4. *若$\mathscr{R}$为代数，则$\mathscr{M}(\mathscr{R}) = \sigma(\mathscr{R})$.*

应用举例：

例2.6.5. *设ξ是概率空间$(\Omega, \mathscr{F}, P)$上的随机变量, F是其分布函数, 则对任意Borel集B, 有*

$$P(\xi \in B) = \int_B dF(x). \tag{2.6.6}$$

这里右边的积分是Lebesgue-Stieltjes积分[2]

这是因为，使上式成立的B的全体构成一单调类（为什么？），且包含了代数：

$$\begin{aligned}\mathscr{E} := \{ \quad & \cup_{i=1}^{n} E_i, n=1,2,\cdots, \text{ 其中} E_i \text{ 或者为}[a,b) -\infty < a \le b \le +\infty, \\ & \text{或者为}(-\infty, b), b \le +\infty\}.\end{aligned}$$

因此包含了$\mathscr{M}(\mathscr{E}) = \sigma(\mathscr{E})$, 而这正是Borel集全体.

定理的证明：

因为$\sigma(\mathcal{R})$为包含$\mathscr{R}$的单调类，而$\mathscr{M}(\mathcal{R})$是包含$\mathscr{R}$的最小的单调类，故

$$\sigma(\mathscr{R}) \supset \mathscr{M}(\mathscr{R}).$$

现在证明相反的包含关系。为此只要证明$\mathscr{M}(\mathscr{R}))$是$\sigma-$代数即可。而这，由引理，只要证明它是代数就行了。

下面就证明它是代数。用命题2.2.9，只要说明它对余及交封闭就行了。先看余。令

$$\mathscr{M} := \{E : E, E^c \in \mathscr{M}(\mathscr{R})\}.$$

自然$\mathscr{M} \supset \mathscr{R}$. 再则易证$\mathscr{M}$是单调类，故$\mathscr{M} \supset \mathscr{M}(\mathscr{R})$.

[2]Henri Lebesgue, 1875-1941, 法国数学家, Borel的学生; Thomas Joannes Stieltjes (1856 - 1894), 荷兰数学家.

再看交。$\forall E \in \mathscr{R}$, 令

$$\mathscr{M}_1(E) := \{F : F \in \mathscr{M}(\mathscr{R}), E \cap F \in \mathscr{M}(\mathscr{R})\}.$$

则$\mathscr{M}(\mathscr{R}) \supset \mathscr{M}_1(E) \supset \mathscr{R}$且$\mathscr{M}_1(E)$为单调类，故$\mathscr{M}_1(E) \supset \mathscr{M}(\mathscr{R})$, 从而

$$\mathscr{M}_1(E) = \mathscr{M}(\mathscr{R}).$$

再对$E \in \mathscr{M}(\mathscr{R}) = \mathscr{M}_1(E)$, 令

$$\mathscr{M}_2(E) := \{F : F \in \mathscr{M}(\mathscr{R}), E \cap F \in \mathscr{M}(\mathscr{R})\}.$$

由上一步, $\mathscr{M}_2(E) \supset \mathscr{R}$, 且也是单调类, 故$\mathscr{M}_2(E) \supset \mathscr{M}(\mathscr{R})$. 这样最后得到

$$\mathscr{M}_2(E) = \mathscr{M}_(\mathscr{R}).$$

于是$\mathscr{M}(\mathscr{R})$是代数. 完毕.

由于在应用中确定一个集类是单调类往往比较简单——我们以后会看到, 它常常有诸如积分的极限定理之类的标准工具可用，因此上面的定理实际上是将确定一个集类为$\sigma-$代数的问题转化为了确定它的一个子类为代数的问题。例如例2.6.5就是如此。但也正是这个例子使我们看到，由于代数的构造仍然比较复杂，真要做起来还是比较烦琐的。这促使E.B.Dynkin对上面传统的单调类定理进行了改造，得到了一种有时候用起来更方便的形式。我们来介绍他的结果。

首先，代替代数的将是π类。

定义2.6.6. 非空集类$\mathscr{C}$若满足

$$E, F \in \mathscr{C} \Longrightarrow E \cap F \in \mathscr{C},$$

则称为π类。

代数当然是π类。π类却不一定是代数。例如

$$\mathscr{C} := \{(-\infty, a), a \in R\}$$

是π类而不是代数。比较一下，它当然比例2.6.5中的代数要简单多了。

代替单调类的将是λ类，它比单调类的要求要多一些——这是可以想到的，因为对代数的要求的减弱必须得到补偿。问题的关键在于，这种多出来的要求在应用中往往很容易就满足了，这样整个问题便得到了简化。

定义2.6.7. 非空集类Λ如果满足以下条件:

1. 包含全空间;

2. $E, F \in \Lambda, E \supset F \Longrightarrow E - F \in \Lambda$;

3. $E_n \in \Lambda, E_n \uparrow \Longrightarrow \lim_n E_n \in \Lambda$,

则称为λ类。

σ代数一定是λ类,但反之不真: λ类对并未必封闭.

同代数等的情形一样，有下面的结果：

定理2.6.8. 设$\mathscr{E}$为非空集类，则存在包含它的最小λ类$\lambda(\mathscr{E})$，称为由$\mathcal{E}$生成的λ类。

现在可以叙述相应的结果了。

引理2.6.9. 一集类为σ代数的充要条件是它既是π类又是λ类。

定理2.6.10. *($\lambda-\pi$定理)若$\mathscr{C}$为π类，则$\lambda(\mathscr{C})=\sigma(\mathscr{C})$.*

我们试一试用现在的定理处理例2.6.5

使(2.6.6)成立的B的全体构成一λ类（为什么？），且包含了π类：

$$\mathcal{C}:=\{(-\infty,b),b\leq+\infty\}.$$

于是结论成立.

看看，果然快得多吧！E.B.Dynkin可真是不寻常啊!

定理的证明: 完全模仿定理2.6.3的证明. 留作练习.

下面我们要叙述的是函数形式的单调类定理. 它与集合形式的对等定理同样有用－－如果不是更有用的话. 不过我们要等到学了积分后才能体会它的妙处.

定理2.6.11. $(X,\mathscr{F})$上的可测函数的子类$\mathscr{H}$若满足

i) $1\in\mathscr{H}$;

ii) $\mathscr{H}$是线性空间;

iii) 若$f_n\geq 0$, $f_n\in\mathscr{H}$, $f_n\uparrow f$, f有界可测(可测), 则$f\in\mathscr{H}$,

iv) $\mathscr{H}\supset\{1_E,E\in\mathscr{C}\}$, 而$\mathscr{C}$是$\pi$类且$\sigma(\mathscr{C})=\mathscr{F}$. 则$\mathscr{H}$包含全体有界可测(可测)函数.

我们会经常碰到满足上面前三条条件的函数类. 为方便计, 我们给它们一个名称.

定义2.6.12. *若函数类$\mathscr{H}$满足上定理中的i)-iii), 则称为$\mathscr{L}$类.*

此定理已可包含大部分常见的应用. 在直接应用它有困难时, 可考虑使用它的一些变化形式－－它们在严加安的《测度论讲义》里有很全面的收集.

证明: 令$\mathscr{G}:=\{E\in\mathscr{F}:1_E\in\mathscr{H}\}$, 则$\mathscr{G}\supset\mathscr{C}$且易证$\mathscr{G}$为$\lambda$类, 于是$\mathscr{G}=\mathscr{F}$. 再由iv)即得定理的结论.

习题:

1. 设X是一空间, $(Y,\mathscr{G})$是一可测空间, f是映射:$X\to Y$. $\forall E\subset Y$, 记

$$f^{-1}(E):=\{x\in X,f(x)\in E\}.$$

证明:
$$f^{-1}(\mathscr{G}) := \{f^{-1}(E) : E \in \mathscr{G}\}$$
为σ代数.

2. 设$\mathscr{E}$为$\mathbb{R}$中的集类, $\sigma(\mathscr{E}) = \mathscr{B}(\overline{\mathbb{R}})$. 设$(X, \mathscr{F})$为测度空间, $f : X \to \overline{\mathbb{R}}$. 证明: f可测的充要条件是:
$$f^{-1}(\mathscr{E}) \subset \mathscr{F}.$$

3. 沿用例2.5.5的记号. 证明一切单调函数都是可测函数.

3 测度

§3.1 代数上的测度

我们已经熟悉了Lebesgue测度. 回忆它的构造: 先在较简单的集合上定义, 然后一步步扩充到较复杂的集合上去. 现在我们要把这种构造方法移植到抽象的框架中去.

定义3.1.1. 设X是全空间, $\mathcal{A}$是其子集代数. $\mathcal{A}$上的非负实值函数μ若满足
$$E, F \in \mathcal{A}, E \cap F = \emptyset \Longrightarrow \mu(E \cup F) = \mu(E) + \mu(F),$$
则称为有限可加集函数; 若满足
$$E_i \in \mathcal{A}, i = 1, 2, \cdots, E_i \cap E_j = \emptyset \quad \forall i \neq j, \quad \cup_{i=1}^{\infty} E_i \in \mathcal{A} \Longrightarrow \mu(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i),$$
则称为有限可列可加集函数, 简称有限测度.

我们先罗列一些简单的性质.

定理3.1.2. *设μ是有限可加集函数, 则它有下列性质:*

1. 可减性: $E, F \in \mathcal{A}, E \subset F \Longrightarrow \mu(F - E) = \mu(F) - \mu(E)$.

2. 单调性: $E, F \in \mathcal{A}, E \subset F \Longrightarrow \mu(E) \leq \mu(F)$.

3. 有限次可加性: $E, E_i \in \mathcal{A}, i = 1, 2, \cdots, n, E \subset \cup_{i=1}^{n} E_i \Longrightarrow \mu(E) \leq \sum_{i=1}^{n} \mu(E_i)$.

4. $E_i \in \mathcal{A}, i = 1, 2, \cdots, E_i \cap E_j = \emptyset \quad \forall i \neq j, \quad \cup_{i=1}^{\infty} E_i \subset E \in \mathcal{A} \Longrightarrow \sum_{i=1}^{\infty} \mu(E_i) \leq \mu(E)$.

若μ是有限测度, 则除此之外还满足:

5. 可列次可加性: $E, E_i \in \mathcal{A}, i = 1, 2, \cdots, E \subset \cup_{i=1}^{n} E_i \Longrightarrow \mu(E) \leq \sum_{i=1}^{\infty} \mu(E_i)$.

证明:

1. 2. 显然.

3. 只需对$n=2$证明, 一般情形可用归纳法. $n=2$时的证明如下:

$$\mu(E)\le\mu(E_1\cup E_2)=\mu(E_1\cup(E_2-E_1))=\mu(E_1)+\mu(E_2-E_1)\le\mu(E_1)+\mu(E_2).$$

4. 只要对每一n证明

$$\sum_{i=1}^{n}\mu(E_i)\le\mu(E).$$

但这是显然的，因为先由有限可加性，次由单调性有

$$\sum_{i=1}^{n}\mu(E_i)=\mu(\cup_{i=1}^{n}E_i)\le\mu(E).$$

5.的证明基本与2.的一样, 只是需要注意由于$\cup_{i=1}^{\infty}E_i$未必属于$\mathcal{A}$, 必须做一点点技术上的修改. 记$F_i:=E\cap E_i$, 则$F_i\in\mathcal{A}$且$\cup_{i=1}^{\infty}F_i=E\in\mathcal{A}$. 于是

$$\begin{aligned}\mu(E) &= \mu(E\cap(\cup_{i=1}^{\infty}E_i))=\mu(\cup_{i=1}^{\infty}F_i)\\ &= \mu(F_1\cup(F_2-F_1)\cup(F_3-F_2-F_1)\cup\cdots)\\ &= \mu(F_1)+\mu(F_2-F_1)+\mu(F_3-F_2-F_1)+\cdots\\ &\le \mu(F_1)+\mu(F_2)+\mu(F_3)+\cdots\\ &\le \mu(E_1)+\mu(E_2)+\mu(E_3)+\cdots\end{aligned}$$

接下来我们要引进集函数的上下连续性概念. 设$\mu:\mathcal{A}\to\mathcal{R}_+$, $E\in\mathcal{R}$. 若$\forall E_n\in\mathcal{A},E_n\downarrow E$, 有$\mu(E_n)\downarrow\mu(E)$, 则称$\mu$在$E$处上连续; 若$\forall E_n\in\mathcal{A},E_n\uparrow E$, 有$\mu(E_n)\uparrow\mu(E)$, 则称$\mu$在$E$处下连续.

现在可以叙述

定理3.1.3. *设μ为$\mathcal{A}$上有限可加集函数, 则下列各条件等价:*

1. μ为测度;

2. μ从下连续;

3. μ从上连续;

4. μ在$\emptyset$处处上连续;

5. μ在X处下连续.

证明：1. $\Longrightarrow$ 2.: 设$E_n,E\in\mathcal{A}$, $,n=1,2,\cdots,E_n\uparrow E$. 由于

$$E=\cup_{n=1}^{\infty}E_n=E_1\cup(E_2-E_1)\cup(E_3-E_2)+\cdots,$$

所以由测度的可列可加性有

$$\begin{aligned}\mu(E) &= \mu(E_1)+\sum_{n=1}^{\infty}\mu(E_{n+1}-E_n)=\lim_{n\to\infty}[\mu(E_1)+\sum_{k=1}^{n}\mu(E_{k+1}-E_k)]\\ &= \lim_{n\to\infty}[\mu(E_1)+\sum_{k=1}^{n}(\mu(E_{k+1})-\mu(E_k))]=\lim_{n\to\infty}\mu(E_n).\end{aligned}$$

2. $\Longrightarrow$ 3.: 设$E_n, E\in\mathcal{A}$, $\quad, n=1,2,\cdots, E_n\downarrow E$. 则$E_n^c, E^c\in\mathcal{A}$, $\quad, n=1,2,\cdots, E_n^c\uparrow E^c$. 于是

$$\mu(E)=\mu(X)-\mu(E^c)=\mu(X)-\lim_{n\to\infty}\mu(E_n^c)=\lim_{n\to\infty}\mu(E_n).$$

3. $\Longrightarrow$ 4. 当然.

4. $\Longrightarrow$ 5. 考虑余集，与2. $\Longrightarrow$ 3.一样.

5. $\Longrightarrow$ 1. 设$E_i\cap E_j=\emptyset, i\neq j$, $E=\cup_{n=1}^{\infty}E_n$. 则

$$\begin{aligned}\mu(E)-\sum_{n=1}^{\infty}\mu(E_n) &= \lim_{n\to\infty}(\mu(E)-\sum_{i=1}^{n}\mu(E_i))\\ &= \lim_{n\to\infty}\mu(E-\cup_{i=1}^{n}E_i)=0.\end{aligned}$$

完毕.

§3.2 测度的可测集

既然σ代数首先是代数, 故上一小节的理论对σ代数全部适用. 但直接在σ代数上构造测度无疑是十分困难的——除去特别简单的场合外，实际上是不可能的. 所以人们一般都是先在代数上构造测度, 再将它扩张到此代数产生的σ代数上去. 这个扩张过程有个放之四海而皆准的一般理论, 不需要个案处理. 但大家要熟悉一些基本的案例, 如Lebesgue测度.

我们先固定记号: X是全空间, $\mathcal{A}$是其子集代数, μ是$\mathcal{A}$上有限测度. 可以很容易地将μ扩张, 因为只要对任一$E\subset X$令

$$\mu^*(E):=\inf\{\sum_{n=1}^{\infty}\mu(A_n), A_n\in\mathcal{A}, E\subset\cup_n A_n\} \tag{3.2.7}$$

就行了. 它称为μ的外测度. 补充$A_0=\emptyset$, 在上式中以$A_n-\cup_{k=0}^{n-1}A_k$代替A_n, 则其测度之和不会变大而其并不会变小, 因而

$$\mu^*(E):=\inf\{\sum_{n=1}^{\infty}\mu(A_n), A_n\in\mathcal{A}, E\subset\cup_n A_n, A_n \text{ 不交}\}. \tag{3.2.8}$$

μ^*自然保持了μ的一些性质, 他们是:

命题3.2.1. *1.* $\mu^*(\emptyset)=0$;

2. $\mu^*(E)\geq 0, \quad \forall E\subset X$;

3. 单调性: $E_1\subset E_2 \Longrightarrow \quad \mu^*(E_1)\leq \mu^*(E_2)$;

4. 强次可加性: $\mu^*(E\cup F)+\mu^*(E\cap F)\leq \mu^*(E)+\mu^*(F), \forall E, F\subset X$;

5. 可列次可加性: $\mu^*(\cup_{n=1}^{\infty}E_n)\leq \sum_{n=1}^{\infty}\mu^*(E_n)$.

证明: 1. 2. 3. 显然. 为证明4., $\forall \varepsilon>0$, 取$\{E_n, F_n\in \mathcal{A}, n=1,2,\cdots\}$及自然数$N$使

$$E\subset \sum_{n=1}^{\infty}E_n, \quad F\subset \sum_{n=1}^{\infty}F_n$$

$$\sum_{n=1}^{\infty}\mu(E_n)\leq \mu^*(E)+\varepsilon\leq \sum_{n=1}^{N}\mu(E_n)+2\varepsilon,$$

$$\sum_{n=1}^{\infty}\mu(F_n)\leq \mu^*(F)+\varepsilon\leq \sum_{n=1}^{N}\mu(F_n)+2\varepsilon.$$

令

$$E'_N:=\sum_{n=1}^{N}E_n, \quad E''_N=\sum_{n=N+1}^{\infty}E_n$$

$$F'_N:=\sum_{n=1}^{N}F_n, \quad F''_N=\sum_{n=N+1}^{\infty}F_n$$

则

$$E\cap F\subset E'_N\cap F'_N+E''_N\cap F'_N+E'_N\cap F''_N+E''_N\cap F''_N$$

$$E\cup F\subset E'_N\cup F'_N\cup E''_N\cup F''_N$$

因而

$$\mu^*(E\cup F)\leq \mu^*(E'_N\cup F'_N)+4\varepsilon$$

$$\mu^*(E\cap F)\leq \mu^*(E'_N\cap F'_N)+6\varepsilon$$

于是

$$\begin{aligned}\mu^*(E\cup F)+\mu^*(E\cap F) &\leq \mu^*(E'_N\cup F'_N)+\mu^*(E'_N\cap F'_N)+10\varepsilon\\ &= \mu^*(E'_N)+\mu^*(F'_N)+10\varepsilon\leq \mu^*(E)+\mu^*(F)+10\varepsilon\end{aligned}$$

但$\varepsilon>0$是任意的, 故结论成立.

最后看5.. $\forall\varepsilon>0$, $\forall n\geq 1$, 找$\{A_{n,m},n,m\geq 1\}$使

$$\mu^*(E_n)\geq\sum_{m=1}^{\infty}\mu(A_{n,m})-2^{-n}\varepsilon.$$

两边对于n求和给出

$$\sum_{n=1}^{\infty}\mu^*(E_n)\geq\sum_{n,m=1}^{\infty}\mu(A_{n,m})-\varepsilon.$$

但$\cup_{n,m=1}^{\infty}A_{n,m}$覆盖了$\cup_{n=1}^{\infty}E_n$, 因此上式右端大于$\mu^*(\cup_{n=1}^{\infty}E_n)$, 这就是说

$$\sum_{n=1}^{\infty}\mu^*(E_n)\geq\mu^*(\cup_{n=1}^{\infty}E_n)-\varepsilon.$$

由ε的任意性便得到5. 完毕.

但很不幸, 可加性丢失了. 恢复可加性需要付出的代价是将μ^*限制在较小的集类上. 这样就引导出可测性的概念.

定义3.2.2. *$E\subset X$若满足$\mu^*(E\cap F)+\mu^*(E^c\cap F)=\mu^*(F)$, $\quad\forall F\subset X$, 则称为μ^*-可测集. 其全体以$\mathscr{F}$表示.*

注意, 由于恒有$\mu^*(E\cap F)+\mu^*(E^c\cap F)\ \geq\ \mu^*(F)$, 故为要$E$是$\mu$-可测集, 只要$\mu^*(E\cap F)+\mu^*(E^c\cap F)\leq\mu^*(F)$ 就够了.

该定义有两个直接的推论. 首先, 它关于"余"是对称的, 因此E和E^c同时为或同时不为μ-可测集. 其次, 设E_1,E_2为μ^*-可测集且不交, 在定义中分别以$E_1,E_1\cup E_2$代E,F, 则得到

$$\mu^*(F\cap(E_1\cup E_2))=\mu^*(E_1\cap F)+\mu^*(E_2\cap F). \tag{3.2.9}$$

进一步取$F=E_1\cup E_2$则得到

$$\mu^*(E_1\cup E_2)=\mu^*(E_1)+\mu^*(E_2), \tag{3.2.10}$$

即μ^*在μ-可测集上保持可加性.

可测集当然是有的.

习题: $\mathscr{A}$的元素全是μ可测集.

但除此而外还有吗?

定理3.2.3. *$\mathscr{F}$为σ代数. 于是特别地, $\mathscr{F}\supset\sigma(\mathscr{A})$.*

由此定理, 可测集是足够地多了.

定理的证明: 由命题2.4.2只要证明$\mathscr{F}$既是代数又对不交的可列并封闭.

我们已经注意到$\mathscr{F}$关于"余"是封闭的.

再设$E_i \in \mathscr{F}$, $i = 1, 2$. 则$\forall F \subset X$, 有

$$\begin{aligned} & \mu^*((E_1 \cup E_2) \cap F) + \mu^*((E_1 \cup E_2)^c \cap F) \\ = \ & \mu^*(E_1 \cap F) + \mu^*(E_1^c \cap E_2 \cap F) + \mu^*(E_1^c \cap E_2^c \cap F) \\ = \ & \mu^*(E_1 \cap F) + \mu^*(E_1^c \cap F) = \mu^*(F). \end{aligned}$$

故$\mathscr{F}$关于并也封闭. 因而$\mathscr{F}$是代数.

现在假设$E_n \in \mathscr{F}, n \geq 1$, $E_i \cap E_i = \emptyset$. 由于$\cup_{k=1}^n E_k, n \geq 1$均是$\mu^*$-可测集, 反复运用3.2.9 便得到, 对任意$F \subset X$有

$$\mu^*(F \cap (\cup_{k=1}^n E_k)) = \sum_{k=1}^n \mu^*(F \cap E_k).$$

这样

$$\begin{aligned} \mu^*(F) &= \mu^*(F \cap (\cup_{k=1}^n E_k)) + \mu^*(F \cap (\cup_{k=1}^n E_k)^c) \\ &\geq \sum_{k=1}^n \mu^*(F \cap E_k) + \mu^*(F \cap (\cup_{n=1}^\infty E_n)^c). \end{aligned}$$

让$n \to \infty$就有

$$\begin{aligned} \mu^*(F) &\geq \sum_{n=1}^\infty \mu^*(F \cap E_k) + \mu^*(F \cap (\cup_{n=1}^\infty E_n)^c) \\ &\geq \mu^*(F \cap (\cup_{n=1}^\infty E_n)) + \mu^*(F \cap (\cup_{n=1}^\infty E_n)^c) \\ &\geq \mu^*(F). \end{aligned} \tag{3.2.11}$$

所以上式中的不等号全为等号,因而$\cup_{n=1}^\infty E_n \in \mathscr{F}$. 完毕.

§3.3 测度的扩张

我们已做好了所有的准备, 可以扩张代数上的测度了.

本小节沿用上小节的记号. 我们的结果是:

定理3.3.1. *μ^*是$\mathscr{F}$上的测度. 另外若ν也是$\mathscr{F}$上的测度且在$\mathscr{A}$上与μ相等, 则在$\sigma(\mathscr{A})$上$\mu^* = \nu$.*

证明: 我们需要证明μ^*在$\mathscr{F}$上的可列可加性. 在(3.2.11)中取$F = \cup_{n=1}^\infty E_n$得

$$\mu^*(\cup_{n=1}^\infty E_n)) = \sum_{n=1}^\infty \mu^*(F \cap E_n) = \sum_{n=1}^\infty \mu^*(E_n).$$

另外, 令

$$\mathscr{G} := \{E \in \sigma(\mathscr{A}) : \mu(E) = \nu(E)\}.$$

则$\mathscr{G}$是单调类且包含代数$\mathscr{A}$, 因而由单调类定理$\mathscr{G} = \mathscr{A}$. 完毕.

有了这一结果, 下面的定义就有意义了——它不会是空集.

定义3.3.2. *测度空间是指一三元组$(X, \mathscr{F}, \mu)$, 其中$(Y, \mathscr{G})$是可测空间, μ是$\mathscr{F}$上测度. 若$\mu(X) = 1$, 则μ称为概率测度, $(X, \mathscr{F}, \mu)$称为概率空间.*

习题: 设$(X, \mathscr{F})$为可测空间, $\mathscr{C}$为π类且$\sigma(\mathscr{C}) = \mathscr{F}$. 若$\mu$, ν为$\mathscr{F}$上的概率测度且在$\mathscr{C}$上有$\mu = \nu$, 则在$\mathscr{F}$上$\mu = \nu$.

我们注意到, 尽管$\mathscr{F}$一般与μ有关——我们以后将写为$\mathscr{F}^\mu$以强调这种关系——但它们都包含了一个公共的σ代数$\sigma(\mathscr{A})$. 于是我们由上一定理立即推出下面漂亮的结果.

定理3.3.3. *$\mathscr{A}$上任意有限测度均可唯一地扩张到$\sigma(\mathscr{A})$上.*

§3.4 Lebesgue-Stieltjes测度

我们考虑前述理论的最基本的例子,即$\mathbb{R}^n$中的Lebesgue-Stieltjes测度. 为简化记号,约定在区间的记法$[a, b)$中,若a为$-\infty$, 则左端点变为开的. 此外, 对不交的并, 将用$\sum$代替$\cup$.

这样,设$X = \mathbb{R}$, 则

$$\mathscr{A} := \{A : A = \sum_{i=1}^{n} [a_i, b_i),\ n \geq 1\}$$

构成代数. 显然, 如果还要求$a_{i-1} < b_{i-1} < a_i < b_i$, 则这种表示法是唯一的.

我们已经说过,这个代数产生的σ代数为$\mathscr{B}^1$.

若μ为$\mathscr{B}^1$上测度, 且$\mu(\mathbb{R}) = 1$, 令

$$F(x) := \mu((-\infty, x)).$$

则我们在初等概率论中已经知道,F满足如下性质:

(i) F单调上升;

(ii) 在任意点$x \in \mathbb{R}$, F左连续,有右极限;

(iii) $F(+\infty) - F(-\infty) = 1$, 其中

$$F(+\infty) := \lim_{x \to +\infty} F(x), \quad F(-\infty) := \lim_{x \to -\infty} F(x).$$

以后我们将满足这三条性质的函数称为*分布函数*. 于是上面的叙述即是说, 任何一个概率测度,都决定着一个分布函数. 现在我们要证明,反过来也是一样,即任何一个分布函数也都决定着一个概率测度. 更准确地,我们有下面的定理:

定理3.4.1. 设F为$\mathbb{R}$上的分布函数,则在$(\mathbb{R},\mathscr{B})$上存在唯一一个概率测度μ使得

$$F(x)=\mu((-\infty,x)).$$

证明: 先证唯一性. 若存在两个测度μ_1, μ_2使得上式成立, 令

$$\mathscr{C}:=\{(-\infty,x),\quad x\in\mathbb{R}\},$$

$$\Lambda:=\{A:\ A\in\mathscr{B},\mu_1(A)=\mu_2(A)\}.$$

则$\mathscr{C}$为π类, Λ为λ类且$\mathscr{C}\subset\Lambda$, 因此由$\lambda-\pi$定理知

$$\Lambda\supset\sigma(\mathscr{C})=\mathscr{B}.$$

为证存在性, 令

$$\mu([a,b)):=F(b)-F(a).$$

而对任意

$$A=\sum_{i=1}^{n}[a_i,b_i),\ n\geq 1\},\quad a_1<b_1<a_2<b_2<\cdots<a_n<b_n$$

令

$$\mu(A):=\sum_{i=1}^{n}\mu([a_i,b_i)).$$

则μ具有如下性质:

设$E,F,E_i,i=1,2,\cdots$为左闭右开的区间, 则

i) $\mu(E)\geq 0,\quad \mu(\emptyset)=0;$

ii) $E\subset F\Longrightarrow\mu(E)\leq\mu(F);$

iii) $E=\cup_{i=1}^{\infty}E_i,\ E_i$ 两两不交 $\Longrightarrow\mu(E)=\sum_{i=1}^{\infty}\mu(E_i)$

i)、ii)是显然的. 下面证明iii).

设$E=[a,b),\ E_i=[a_i,b_i)$. $\forall n$ 我们有

$$\cup_{i=1}^{n}E_i\subset E.$$

必要时将$E_i,i=1,\cdots,n$重新排序并删除掉空集, 可假设$a_1<b_1\leq a_2<\cdots\leq a_n<b_n$. 于是

$$\mu(E)=F(b)-F(a)\geq F(b_n)-F(a_1)\geq\sum_{i=1}^{n}(F(b_i)-F(a_i))=\sum_{i=1}^{n}\mu(E_i).$$

令$n\to\infty$得

$$\mu(E)\geq\sum_{i=1}^{\infty}\mu(E_i).$$

现在证明反方向的不等式. 不妨设$a < b$.

$\forall \varepsilon > 0$. 由F的左连续性知存在$\exists b' \in (a,b)$使

$$F(b) - F(b') < \varepsilon.$$

又$\forall n$, $\exists a'_n$使

$$F(a_n) - F(a'_n) < \frac{\varepsilon}{2^n}.$$

由于$\cup_{i=1}^{\infty}(a'_n, b_n) \supset [a, b']$, 故$\exists n$使$\cup_{i=1}^{n}(a'_i, b_i) \supset [a, b']$. 因而

$$F(b') - F(a) \leq \sum_{i=1}^{n}(F(b_i) - F(a'_i)).$$

这样,

$$\sum_{i=1}^{\infty}\mu(E_i) \geq \sum_{i=1}^{n}\mu(E_i) \geq \sum_{i=1}^{n}(F(b_i) - F(a'_i)) - \varepsilon \geq \mu(E) - 2\varepsilon.$$

由ε的任意性便有

$$\sum_{i=1}^{\infty}\mu(E_i) > \mu(E).$$

现在证明μ是可列可加的. 若$E, E_i \in \mathscr{A}$, $E = \cup_{i=1}^{\infty}E_i$, E_i 两两不交. 设$E_i = \cup_{j=1}^{n_i}E_{ij}$, $E = \cup_{j=1}^{m}F_j$, E_{ij}, $F_j \in \mathcal{P}$ 则由于

$$F_j = E \cap F_j = \cup_{i=1}^{\infty} \cup_{k=1}^{n_i} (E_{ik} \cap F_j),$$

我们有

$$\mu(F_j) = \sum_{i=1}^{\infty}\sum_{k=1}^{n_i}\mu(E_{ik} \cap F_j).$$

所以

$$\begin{aligned}
\mu(E) &= \sum_{j=1}^{m}\mu(F_j) = \sum_{j=1}^{m}\sum_{i=1}^{\infty}\sum_{k=1}^{n_i}\mu(E_{ik} \cap F_j) \\
&= \sum_{i=1}^{\infty}\sum_{k=1}^{n_i}\sum_{j=1}^{m}\mu(E_{ik} \cap F_j) \\
&= \sum_{i=1}^{\infty}\sum_{k=1}^{n_i}\mu(\cup_{j=1}^{m}(E_{ik} \cap F_j)) \\
&= \sum_{i=1}^{\infty}\sum_{k=1}^{n_i}\mu(E_{ik}) \\
&= \sum_{i=1}^{\infty}\mu(E_i).
\end{aligned}$$

于是由定理3.3.3, 我们立即知道μ可唯一地扩张到$\mathscr{B}$上成为有限测度.

这样构造出来的测度就称为相应于F的Lebesgue-Stieltjes测度. 为表示μ和F的关系, 我们常常将μ记为μ_F.

例3.4.2. 设

$$F(x) = \begin{cases} 0, \ x < 0 \\ x, \ x \in [0,1] \\ 1, \ x > 1 \end{cases} \tag{3.4.12}$$

这时F对应的测度即为$[0,1]$上的*Lebesgue*测度.

例3.4.3. 设$a_i > 0, \ i = 1, 2, \cdots, \ \sum_{i=1}^{\infty} a_i = 1, \ x_i \in \mathbb{R}$. 令

$$F(x) := \sum_{x_i \le x} a_i.$$

则对应的概率测度为离散分布,即

$$\mu(\{x_i\}) = a_i.$$

现在我们来看高维空间即$\mathbb{R}^n$的情形.

设F是定义在$\mathbb{R}$上的函数, 记$x = (x_1, \cdots, x_n)$. 假定

(i) F左连续,即

$$\lim_{x^{(k)} \uparrow x} F(x^{(k)}) = F(x),$$

其中$x^{(k)} \uparrow x$表示对任意$1 \le i \le n$, $x_n^{(k)} \uparrow x_n$;

(ii)对任意$a = (a_1, \cdots, a_n) \le b = (b_1, \cdots, b_n) \in \mathbb{R}$,

$$\Delta_{ab} F \ge 0,$$

其中

$$\Delta_{ab} := \Delta_{a_i b_i} \cdots \Delta_{a_n b_n},$$

而

$$\Delta_{a_i b_i} F(x_1, \cdots, x_n) = F(x_1, \cdots, x_{i-1}, b_i, x_{i+1}, \cdots, x_n) - F(x_1, \cdots, x_{i-1}, a_i, x_{i+1}, \cdots, x_n).$$

则有

定理3.4.4. 在$(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$上存在唯一一个测度$\mu_F$满足

$$\mu_F([a, b)) = \Delta_{ab} F.$$

例3.4.5. 设

$$F(x) = \begin{cases} 0, \ 若\exists i, \ s.t. \ x_i < 0 \\ x_1 \cdots x_n, \ x \in [0,1]^n \\ 1, \ 若\exists i, \ s.t. \ x_i > 1 \end{cases} \tag{3.4.13}$$

这时F对应的测度即为$[0,1]^n$上的*Lebesgue*测度.

§3.5 不可测集

有了可测集的概念后,自然要问的一个问题是:是否存在不可测集?答案是肯定的.下面就是一个简单例子.

对于$x \in \mathbb{R}$, 令$[x]$表示其整数部分, $\{x\}$表示其小数部分. 即$[x]$为不超过x的最大整数, $\{x\} = x - [x]$. 设α为一无理数.

考虑$[0,1)$上的Lebesgue测度. 对$x, y \in [0,1)$, 定义等价关系:

$$x \sim y \Longleftrightarrow \exists n \in \mathbb{N} \text{ s.t. } \{x + n\alpha\} = y.$$

若$x \sim y$, 则把x, y归为一类. 于是$[0,1)$被分成了若干等价类. 在每一等价类中取一个元素,构成集合E. 对任意$n \in \mathbb{N}$, 令

$$E_n := \{y : \{y + n\alpha\} \in E\}.$$

则$\{E_n\}$两两不交. 若E可测, 则由Lebesgue测度的平移不变性有

$$\mu(E_n) = \mu(E), \quad \forall n \in \mathbb{N}.$$

但

$$\mu([0,1)) = \sum_{n=-\infty}^{\infty} \mu(E_n),$$

故$\mu(E) = 0$时,上式为零; $\mu(E) > 0$时,上式为无穷大. 这两者皆不可能,因此E不可测.

§3.6 测度的完备化

μ在$\sigma(\mathscr{A})$上的唯一扩张μ^*以后仍将记为μ. 那么$\mathscr{F}^\mu$与$\sigma(\mathscr{A})$的关系如何呢？要说明此事，需要用到完备测度的概念.

定义3.6.1. *设$(Y, \mathscr{G}, \nu)$为测度空间. 若*

$$E \in \mathscr{G}, \nu(E) = 0 \Longrightarrow \quad \forall F \subset E, F \in \mathscr{G}.$$

则ν称为完备测度, $(Y, \mathscr{G}, \nu)$称为完备测度空间.

测度空间当然不会自动成为完备测度空间, 但我们有办法强迫它成为.

定理3.6.2. *设$(Y, \mathscr{G}, \nu)$为测度空间. 令*

$$\begin{aligned} \bar{\mathscr{G}}^\nu : &= \{E \cup N, E \in \mathscr{G}, \quad , \exists F \in \mathscr{G}, \quad \nu(F) = 0 \text{ 使} N \subset F\} \\ &= \{A \subset Y ; \exists E_1, E_2 \in \mathscr{G}, \quad \nu(E_1) = \nu(E_2) \text{ 使} E_1 \subset A \subset E_2\}; \end{aligned}$$

$$\bar{\nu}(E \cup N) := \nu(E).$$

则$(Y, \bar{\mathscr{G}}^\nu, \bar{\nu})$为完备测度空间.

证明很容易, 读者自行之.

$(Y, \bar{\mathscr{G}}^{\nu}, \bar{\nu})$称为$(Y, \mathscr{G}, \nu)$的完备化.

我们有:

定理3.6.3. $(X, \mathscr{F}^{\mu}, \mu)$为$(X, \sigma(\mathcal{A}), \mu)$的完备化.

证明: 首先证明

$$\mu^*(E) = \inf\{\mu(F) : E \subset F, F \in \sigma(\mathcal{A})\},$$

并且此下确界可以达到.

事实上, 按定义有

$$\begin{aligned} \mu^*(E) &= \inf\{\sum_{n=1}^{\infty} \mu(E_n), E_n \in \mathcal{A}, E \subset \cup_{n=1}^{\infty} E_n\} \\ &\geq \inf\{\mu(\cup_{n=1}^{\infty}(E_n)), E_n \in \mathcal{A}, E \subset \cup_{n=1}^{\infty} E_n\} \\ &\geq \inf\{\mu(F) : E \subset F, F \in \sigma(\mathcal{A})\} \\ &\geq \mu^*(E), \end{aligned}$$

因而其中的三个不等式全为等式. 再取$F_n \in \mathscr{F}$, $E \subset$使$\mu(F_n) \leq \mu^*(E) + \frac{1}{n}$, 则$F := \cap_{n=1}^{\infty} F_n \in \mathscr{F}$, $E \subset F$ 且$\mu^*(E) = \mu(F)$.

设$E \in \mathscr{F}^{\mu}$, 则$E^c \in \mathscr{F}$. 取$F_1, F_2 \in \mathscr{F}$, $E_1 \subset F$, $E^c \subset F_2$, $\mu(F_1) = \mu^*(E)$, $\mu(F_2) = \mu^*(E^c)$, 则$F_2^c \subset E \subset F_1$ 且

$$\mu(F_2^c) = \mu(X) - \mu(F_2) = \mu(X) - \mu^*(E^c) = \mu^*(E) = \mu(F_1).$$

完毕.

§3.7 一般测度

至今我们一直假定所考虑的测度是有限的，即$\mu(X) < +\infty$. 应用中常常也会碰到一类比此稍微广泛些的测度，它只满足下面较弱的条件：

存在$A_n \in \mathscr{A}, n \geq 1$使$X = \cup_{n=1}^{\infty} A_n$而$\mu(A_n) < +\infty \quad \forall n$.

这种测度叫σ-有限测度. 对这样的测度, 我们总可以先将它限制在每个$(X, A_n \cap \mathcal{A})$上, 利用对有限测度已有的结论进行扩张, 然后再合并起来, 得到$(X, \sigma(\mathcal{A}))$的σ-有限测度.

例3.7.1. 设

$$F(x) = x, \quad x \in \mathbb{R}.$$

则对每一$n\mathbb{N}$, 在$[n, n+1)$上F都产生一个*Lebesgue*测度. 把这些测度拼起来则得到$\mathbb{R}$上的*Lebesgue*测度. 高维情况可同样处理.

例3.7.2. 一般地,可以考虑任何单调上升的左连续函数F,而不必假设

$$F(+\infty)-F(-\infty)=1.$$

这时,同上面一样,只需先在每个$[n,n+1)$上产生*Lebesgue-Stieltjes*测度,然后把各段的测度拼起来即可.

σ有限测度有和有限测度类似的连续性性质，即

命题3.7.3. 设$(X,\mathscr{F},\mu)$为σ有限测度空间,则

$$\{E_n\}\subset\mathscr{F},\ E_n\uparrow E\Longrightarrow\mu(E_n)\uparrow\mu(E)$$

$$\{E_n\}\subset\mathscr{F},\ E_n\downarrow E,\mu(A_1)<\infty\Longrightarrow\mu(E_n)\downarrow\mu(E)$$

注意第二个结论的条件$\mu(A_1)<\infty$, 这是必须的, 因为例如假设

$$X=\mathbb{R},\quad E_n=(n,\infty),\quad E=\emptyset,\quad \mu=\ \text{Lebesgue测度}$$

则$\mu(E)=0$而$\mu(E_n)=\infty,\ \forall n$.

推论3.7.4. 设$(X,\mathscr{F},\mu)$为σ有限测度空间, $\{E_n\}\subset\mathscr{F}$. 则

$$\mu(\liminf_{n\to\infty}E_n)\le\liminf_{n\to\infty}\mu(E_n).$$

$$\mu(\cup_{n=1}^{\infty}E_n)<\infty\Longrightarrow\mu(\limsup_{n\to\infty}E_n)\ge\limsup_{n\to\infty}\mu(E_n).$$

证明只是按上、下极限的定义直接应用上命题.

最后我们指出,如果连σ有限也不要求,就得到一般的测度. 上述性质对一般的测度也是成立的. 当然一般的测度在应用中较少碰到.

§3.8 测度空间上的可测函数

本节设$(X,\mathscr{F},\mu)$为完备测度空间. 此时, 由于$(X,\mathscr{G})$是可测空间, 因而可测函数是已经有定义了的. 但因为现在多了一个测度, 所以研究的内容会更丰富一些: 这主要表现在有了两种不同的收敛性.

首先, 我们约定一个涉及X中的元素的命题如果除开一个零测集——即测度为零的集合——外处处成立, 则叫它几乎处处成立, 并用a.e.——almost everywhere——表示.

定义3.8.1. 设$\{f_n,n\ge 1\}$, f为可测函数. 若$\mu(\{x\in X:f_n(x)\nrightarrow f(x)\})=0$, 则称$f_n$几乎处处收敛到$f$, 记为$f_n\to f$ *a.e.*.

若$f(x)$有限, 用$\varepsilon - N$语言，使$f_n(x)$收敛到$f(x)$的集合可表为

$$\begin{aligned} & \{x \in X : \forall \varepsilon > 0, \exists N, \text{ 当} n \geq N\text{时}, |f_n(x) - f(x)| < \varepsilon\} \\ = & \cap_{\varepsilon>0} \cup_{N=1}^{\infty} \cap_{n=N}^{\infty} \{x \in X : |f_n(x) - f(x)| < \varepsilon\} \\ = & \cap_{\varepsilon_k>0} \cup_{N=1}^{\infty} \cap_{n=N}^{\infty} \{x \in X : |f_n(x) - f(x)| < \varepsilon_k\} \end{aligned}$$

其中$\{\varepsilon_n, n \geq 1\}$是任意一列单调下降到0的常数. 因此

$$\begin{aligned} & \{x \in X : f_n \nrightarrow f\} \\ = & \cup_{\varepsilon_k>0} \cap_{N=1}^{\infty} \cup_{n=N}^{\infty} \{x \in X : |f_n(x) - f(x)| \geq \varepsilon_k\} \end{aligned}$$

注意$\cap_{N=1}^{\infty} \cup_{n=N}^{\infty} \{x \in X : |f_n(x) - f(x)| \geq \varepsilon_k\}$随$k$上升而上升, 于是

$$\mu(\{x \in X : f_n(x) \nrightarrow f(x)\}) = \lim_{k\to\infty} \mu(\cap_{N=1}^{\infty} \cup_{n=N}^{\infty} \{x \in X : |f_n(x) - f(x)| \geq \varepsilon_k\}),$$

其左边为零当且仅当右边每一项为零. 这样得到

命题3.8.2. *若f_n几乎处处有限, f_n几乎处处收敛到f的充要条件是:*

$$\mu(\cap_{N=1}^{\infty} \cup_{n=N}^{\infty} \{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}) = 0, \qquad \forall \varepsilon > 0. \tag{3.8.14}$$

若我们事先并不知道极限函数f, 只是为了判别f_n是否几乎处处收敛到一几乎处处有限的可测函数, 就可以利用数列收敛的Cauchy准则而得到其收敛集合为

$$\begin{aligned} & \{x \in X : \forall \varepsilon > 0, \exists N, \text{ 当} m, n \geq N\text{时}, |f_m(x) - f_n(x)| < \varepsilon\} \\ = & \cap_{\varepsilon>0} \cup_{N=1}^{\infty} \cap_{m,n=N}^{\infty} \{x \in X : |f_m(x) - f_n(x)| < \varepsilon\} \\ = & \cap_{\varepsilon_k>0} \cup_{N=1}^{\infty} \cap_{m,n=N}^{\infty} \{x \in X : |f_m(x) - f_n(x)| < \varepsilon_k\} \end{aligned}$$

由此出发, 与上一命题类似地, 可以证明

命题3.8.3. *若f_n几乎处处有限, f_n几乎处处收敛到一几乎处处有限的可测函数的充要条件是:*

$$\mu(\cap_{N=1}^{\infty} \cup_{m,n=N}^{\infty} \{x \in X : |f_m(x) - f_n(x)| \geq \varepsilon\}) = 0, \qquad \forall \varepsilon > 0.$$

我们再给出一个充分条件. 补充$f_0 = 0$后, 自然可写

$$f_n = \sum_{k=1}^{n} (f_k - f_{k-1}).$$

于是f_n收敛相当于右边的级数收敛, 而这又只要它绝对收敛. 因此

$$\{x \in X : \ f_n\text{收敛}\} \supset \{x \in X : \sum_{n=1}^{\infty} |f_n(x) - f_{n-1}(x))| < \infty\}.$$

也就是说

$$\{x\in X:\ f_n\text{不收敛}\}\subset\{x\in X:\sum_{n=1}^{\infty}|f_n(x)-f_{n-1}(x))|=\infty\}.$$

现设$\{\varepsilon_n,n>1\}$使$\sum_{n=1}^{\infty}\varepsilon_n<\infty$, 则

$$\cup_{n=1}^{\infty}\cap_{k=n}^{\infty}\{x\in X:|f_k(x)-f_{k-1}(x)|<\varepsilon_k\}\subset\{x\in X:\sum_{n=1}^{\infty}|f_n(x)-f_{n-1}(x))|<\infty\}.$$

因而

$$\{x\in X:\sum_{n=1}^{\infty}|f_n(x)-f_{n-1}(x))|=\infty\}\subset\cap_{n=1}^{\infty}\cup_{k=n}^{\infty}\{x\in X:|f_k(x)-f_{k-1}(x)|<\varepsilon_k\}.$$

所以

$$\begin{aligned}
&\mu(\{x\in X:\sum_{n=1}^{\infty}|f_n(x)-f_{n-1}(x))|=\infty\})\\
\le\ &\mu(\cap_{n=1}^{\infty}\cup_{k=n}^{\infty}\{x\in X:|f_k(x)-f_{k-1}(x)|\ge\varepsilon_n\})\\
=\ &\lim_{n\to\infty}\mu(\cup_{k=n}^{\infty}\{x\in X:|f_k(x)-f_{k-1}(x)|\ge\varepsilon_k\})\\
\le\ &\lim_{n\to\infty}\sum_{k=n}^{\infty}\mu(\{x\in X:|f_k(x)-f_{k-1}(x)|\ge\varepsilon_k\}).
\end{aligned}$$

其最后一项为零的充要条件是

$$\sum_{n=1}^{\infty}P(\{x\in X:|f_n(x)-f_{n-1}(x)|\ge\varepsilon_n\})<\infty.$$

这样, 我们得到

命题3.8.4. 若正数列$\{\varepsilon_n,n\ge1\}$使$\sum_n\varepsilon_n<\infty$, 而$f_n$又满足

$$\sum_{n=1}^{\infty}P(\{x\in X:|f_n(x)-f_{n-1}(x)|\ge\varepsilon_n\})<\infty,$$

则f_n几乎处处收敛到一几乎处处有限的可测函数.

什么样的序列满足此命题的条件呢?
若f_n为依测度Cauchy列, 即

$$\lim_{m,n\to\infty}\mu(\{|f_m(x)-f_n(x)|\ge\varepsilon\})=0,\forall\varepsilon>0.$$

取n_1使$m, n \geq n_1$时

$$\mu(|f_m - f_n| \geq 2^{-1}) < 2^{-1}.$$

一般地, 对$j \geq 2$, 设已取好n_{j-1}, 则取$n_j > n_{j-1}$使$m, n \geq n_j$时有

$$\mu(|f_m - f_n| \geq 2^{-j}) < 2^{-j}.$$

这样

$$\mu(|f_{n_{j+1}} - f_{n_j}| \geq 2^{-j}) < 2^{-j}.$$

于是

$$\sum_{j=1}^{\infty} \mu(|f_{n_{j+1}} - f_{n_j}| \geq 2^{-j}) < \infty,$$

也就是说对$\varepsilon_j = 2^{-j}$, f_{n_j}满足上命题的条件. 因此有

推论3.8.5. *若f_n为依测度Cauchy列, 则在其中可抽出一子列几乎处处收敛到一几乎处处有限的随机变量.*

我们再回过头来再看3.8.14. 由于$\cup_{k\geq n}\{|f_k - f| > \varepsilon\}$随$n$的上升而下降，故3.8.14相当于

$$\lim_{n\to\infty} \mu(\cup_{k\geq n}\{|f_k - f| > \varepsilon\}) = 0.$$

将此条件减弱, 得到

定义3.8.6. *设$\{f_n, n \geq 1\}$, f为可测函数. 若$\forall \varepsilon > 0$, $\lim_{n\to\infty} \mu(\{|f_n(x) - f(x)| \geq \varepsilon\}) = 0$, 则称$f_n$依测度收敛到$f$, 记为$f_n \to f(\mu)$.*

注意, 由定义, 几乎处处收敛到某几乎处处有限的可测函数自然蕴涵着依测度收敛到同一函数. 这里"几乎处处有限"六字是必须的: 若f在一正测度的集合上为∞, 则不可能符合定义.

另外, 我们有如下的定理:

定理3.8.7. *f_n依测度收敛的充要条件是f_n为依测度Cauchy列.*

证明: 必要性由不等式

$$\mu(|f_n - f_m| > \varepsilon) \leq \mu(|f_n - f| > \varepsilon/2) + \mu(|f_m - f| > \varepsilon/2)$$

看出. 现在证充分性. 首先, 用推论3.8.5, 在$\{f_n, n \geq 1\}$中取子列$\{f_{n_j}, j \geq 1\}$几乎处处收敛. 设极限函数是f. 在不等式

$$\mu(|f_n - f| > \varepsilon) \leq \mu(|f_n - f_{n_j}| > \varepsilon/2) + \mu(|f_{n_j} - f| > \varepsilon/2)$$

中让n, n_j都趋于无穷大即得结论. 完毕.

由此定理及推论3.8.5我们又有

推论3.8.8. *依测度收敛的序列中必可抽出子列几乎处处收敛.*

4 积分

一般测度空间上积分的定义和$\mathbb{R}$上的Lebesgue积分是完全一样的，因此熟悉大学实变函数课程的人再来学习这一理论是易如连掌都不用反的.

我们固定一个有限测度空间$(X, \mathscr{F}, \mu)$. 我们从简单情形开始.

§4.1 简单可测函数的积分

定义4.1.1. *设f是简单可测函数:*

$$f = \sum_{i=1}^{n} a_i 1_{E_i},$$

其中诸E_i不交. f的积分定义为

$$\int f d\mu = \sum_{i=1}^{n} a_i \mu(E_i).$$

这个定义是合理的, 也就是说, 若f还有另一种表达式, 那么计算出来的值不变. 事实上, 设另有

$$f = \sum_{j=1}^{m} b_j 1_{F_j}, \quad b_j \geq 0,$$

这里诸F_j也是不交的. 拿掉其中为零的后, 可以进一步假定诸a_i, b_j均不为零, 这样势必有

$$\cup_{i=1}^{n} E_i = \cup_{j=1}^{m} F_j.$$

于是

$$E_i = \cup_{j=1}^{m} (E_i \cap E_j), \quad \forall\ i.$$

因此, 由测度的可加性有

$$\sum_{i=1}^{n} a_i \mu(E_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \mu(E_i \cap E_j).$$

但除非$E_i \cap F_j = \emptyset$, 否则必有$a_i = b_j$, 因此上式右端又等于(再用一次测度的可加性)

$$\sum_{j=1}^{m} \sum_{i=1}^{n} b_j \mu(E_i \cap E_j) = \sum_{j=1}^{m} b_j \mu(F_j).$$

因此按不同出表达式定义出来的积分值是相同的.

而在定义中所要求的不交假设也可以去掉.

命题4.1.2. *设f是简单可测函数:*

$$f=\sum_{i=1}^{n}a_i1_{E_i},\quad a_i\geq 0,$$

则

$$\int fd\mu=\sum_{i=1}^{n}a_i\mu(E_i).$$

证明:我们利用2.5.4

$$f=\sum_{A}(\sum_{i\in A}a_i)1_{F_A}.$$

于是, 注意到$\cup_{A\ni i}F_A=E_i$, 就有

$$\int fd\mu \;=\; \sum_{A}(\sum_{i\in A}a_i)\mu(F_A)=\sum_{i=1}^{n}a_i(\sum_{A\ni i}\mu(F_A))=\sum_{i=1}^{n}a_i\mu(E_i).$$

完毕.

简单可测函数的积分有如下性质:

命题4.1.3. *1) 非负性: $f\geq 0\Longrightarrow\int f\delta\mu\geq 0$;*

2) 线性性: $\int(af+bg)d\mu=a\int fd\mu+b\int gd\mu$;

3) 单调性: $f\geq g$ a.e. $\Longrightarrow\int f\ d\mu\geq\int g\ d\mu$;

4) $|\int\ fd\mu<\int|f|\ d\mu$.

5) $f=g$ a.e. $\Longrightarrow\int f=\int g$.

6) $\forall f$, $\mu(E)=0\Longrightarrow\int_E fd\mu=0$.

7) Chebyshev不等式: $\mu(|f|>c)\leq\frac{1}{c}\int|f|d\mu$, $\forall c>0$.

证明: 6)由5)推出,4) 5) 7)由3)推出, 3)由1)和2)推出, 2)直接由定义推出. 1)也很简单: 若f表为

$$f=\sum_{i=1}^{n}a_i1_{E_i},$$

其中诸E_i不交, 则f非负时每个a_i必非负, 再直接由定义即得. 完毕.

§4.2 有界可测函数的积分

设f是有界可测函数. 由命题2.5.9的证明, 有简单函数列f_n, $n\geq 1$, 一致收敛到f. 于是我们有

$$|\int f_n\ d\mu-\int f_m\ d\mu|\leq\int|f_n-f_m|d\mu\leq(\sup_{x\in X}|f_n(x)-f_m(x)|)\mu(X),$$

从而$\int f_nd\mu$是Cauchy列. 因而极限存在. 我们由此获得

定义4.2.1. *有界可测函数f的积分定义为*

$$\int f d\mu := \lim_n \int f_n d\mu,$$

其中f_n是任意一致收敛到f的简单可测函数列.

我们仍然必须说明此定义的合理性. 设f_n, g_n是两列非负简单可测函数, 均一致收敛到f. 令

$$h_{2n+1} = f_n, \quad h_{2n} = g_n,$$

则h_n也是一致收敛到f的非负简单可测函数列, 因而$\lim_n \int h_n d\mu$存在, 故$\lim_n \int f_n d\mu$与$\lim_n \int g_n d\mu$一定相等.

用极限过渡, 容易证明

命题4.2.2. *命题4.1.3的结论对有界可测函数照样成立.*

§4.3 非负可测函数的积分

对未必有界的非负可测函数f, 令

$$T_n f(x) := \begin{cases} f(x) & \text{若} f(x) \le n, \\ n, & \text{若} f(x) \ge n. \end{cases}$$

则$T_n f \uparrow f$, 从而$\int T_n f \ d\mu \uparrow$. 我们定义

定义4.3.1.

$$\int f d\mu := \lim_n \int T_n f \ d\mu.$$

用极限过渡,容易证明

命题4.3.2. *命题4.2.2的结论对非负可测函数照样成立.*

§4.4 可测函数的积分

对一般的可测函数f, 定义

定义4.4.1. *若$\int f^+ \ d\mu$与$\int f^- \ d\mu$中至少一个不为∞, 则称其积分存在且其值为*

$$\int f \ d\mu := \int f^+ \ d\mu - \int f^- \ d\mu;$$

若两个均$< \infty$, 则称其可积.

当有必要明确积分的范围时, 我们就用$\int_X f \ d\mu$代替$\int f \ d\mu$

我们还是有

命题4.4.2. *命题4.2.2的结论对可测函数照样成立.*

另外, 对任一可测集$A \in \mathscr{F}$, 定义在A上的积分为

$$\int_A f d\mu := \int f 1_A d\mu.$$

§4.5 σ有限测度空间上的积分

现设$(X, \mathscr{F}, \mu)$为σ有限测度空间: 即存在$\{X_n \in \mathscr{F}, n \geq 1\}$使$X_n \uparrow X$且$\mu(X_n) < \infty$. 我们要在此空间上定义积分.

首先对非负有界的可测函数定义. 设$0 \leq f \leq M$, M为非负常数. 定义

$$\int f\ d\mu := \lim_n \int_{X_n} f\ d\mu.$$

像过去一样, 我们必须说明此定义的合理性. 设另有一列$\{Y_n \in \mathscr{F}, n \geq 1\}$使$Y_n \uparrow X$且$\mu(Y_n) < \infty$. 固定任一自然数$k$, 由于$X_n \cap Y_k \uparrow Y_k (n \uparrow \infty)$, 所以$\mu(X_n \cap Y_k) \uparrow \mu(Y_k)$. 于是

$$\begin{aligned}
& \int_{Y_k} f\ d\mu - \int_{X_n} f\ d\mu \\
\leq\ & \int_{Y_k} f\ d\mu - \int_{X_n \cap Y_k} f\ d\mu \\
\leq\ & \int_{Y_k} f(1 - 1_{X_n \cap Y_k})\ d\mu \\
\leq\ & M(\mu(Y_k) - \mu(Y_k \cap X_n)) \to 0, \quad n \to \infty.
\end{aligned}$$

因而

$$\int_{Y_k} f\ d\mu \leq \lim_n \int_{X_n} f\ d\mu,$$

这意味着

$$\lim_n \int_{Y_n} f\ d\mu \leq \lim_n \int_{X_n} f\ d\mu.$$

同理

$$\lim_n \int_{X_n} f\ d\mu \leq \lim_n \int_{Y_n} f\ d\mu.$$

于是

$$\lim_n \int_{Y_n} f\ d\mu = \lim_n \int_{X_n} f\ d\mu.$$

所以定义是合理的.

其次, 对非负可测函数f, 跟有限测度的情形一样, 定义

$$\int f d\mu := \lim_n \int T_n f d\mu.$$

最后, 对一般的可测函数f, 若$\int f^+ d\mu$与$\int f^- d\mu$中之一有限, 则说$\int f d\mu$存在且定义

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu;$$

而若两个均有限, 则说f可积.

对任意$A \in \mathscr{F}$, 定义A上的积分为

$$\int_A f d\mu := \int f 1_A d\mu.$$

至此, 我们定义积分的任务已全部完成. 通过若干次中间步骤过渡, 可以证明:

命题4.5.1. *命题4.2.2的结论照样成立.*

除此而外, 我们还有:

命题4.5.2. *1) f可积$\iff$ $|f|$可积;*
2) f可积$\implies \mu(|f| = \infty) = 0$.

其中第一式直接由定义, 第二式可如下证明: 由Chebyshev不等式, 对任意自然数n有

$$\mu(|f| \geq n) \leq n^{-1} \int f d\mu,$$

因此

$$\mu(|f| = \infty) = \lim_n \mu(|f| \geq n) = 0.$$

习题1: 设f有界, 若对任意$A \in \mathscr{B}$,

$$\int_A f d\mu = 0,$$

则$f = 0$ a.e..

回忆关于级数收敛的下列结果: 若$a_n \geq 0$且

$$\sum_{n=1}^{\infty} a_n < \infty,$$

则存在b_n满足

$$\lim_n \frac{b_n}{a_n} = \infty$$

且

$$\sum_n b_n < \infty.$$

对于积分也有类似结果.

命题4.5.3. *设X为可积随机变量,则存在增凸函数$h : \mathbb{R}_+ \to \mathbb{R}_+$, $\lim_{x\to\infty} \frac{h(x)}{x} = \infty$使$h(|X|)$可积.*

Proof. 因为
$$\lim_{n\to\infty} E[|X|1_{\{|X|\geq n\}}]=0,$$
所以存在$n_k\uparrow\infty$使$n_0=0$
$$E[|X|1_{\{|X|\geq n_k\}}]\leq 2^{-k},\quad \forall k>0,$$
且$\frac{n_{k+1}}{n_{k+1}-n_k}<2$. 于是$k+\frac{n_{k+1}}{n_{k+1}-n_k}$非降。

令$h(x)$为在每个$[n_k,n_{k+1}]$上为线性的函数且
$$h(n_k)=kn_k.$$
则h满足命题的条件且在$[0,n_k]$上$h(x)\leq kx$. 故
$$E[h(|X|)]\leq\sum_k k2^{-k}<\infty.$$
□

从这个意义上讲,永远不存在一个"最大的"可积随机变量.

5 积分号下取极限

§5.1 有限测度空间情形

固定着有限测度空间$(X,\mathscr{F},\mu)$, 我们来看如何在积分号下取极限. 我们不久就会看到, 相对于Riemann积分, 现在进行这种运算的环境要宽松得多.

首先有

定理5.1.1. *(有界收敛定理)若$f_n\to f(\mu)$, 且存在$M>0$使$|f_n|\vee|f|\leq M$, 则*
$$\lim_n\int f_n d\mu=\int f d\mu.$$

证明: $\forall\varepsilon>0$有
$$\begin{aligned}\int|f-f_n|d\mu &= \int_{|f-f_n|>\varepsilon}|f-f_n|d\mu+\int_{|f-f_n|<e}|f-f_n|d\mu\\ &\leq M\mu(|f-f_n|>\varepsilon)+\varepsilon\mu(X).\end{aligned}$$
先令$n\to\infty$次令$\varepsilon\to0$便得结论.

注意由于在$\mu(X)<\infty$时, 几乎处处收敛蕴涵依测度收敛, 故上定理在几乎处处收敛时也成立.

下面是Fatou-Beppo Levi的单调收敛定理.

定理5.1.2. *设$\{f_n\}$是上升的非负可测函数列. 令$\lim_n f_n = f$ a.e. . 则f的积分存在且*

$$\lim_n \int f_n d\mu = \int f d\mu.$$

证明: 因为$\int f_n d\mu \le \int f d\mu$, 故若$\int f_n d\mu \uparrow \infty$, 则必有$\int f d\mu = \infty$.

现假设存在$C > 0$使$\int f_n d\mu \le C$. 因为$\forall M > 0$有$T_M(f_n) \to T_M(f)$ a.e., 故由上定理有

$$\lim_n \int T_M f_n d\mu = \int T_M f d\mu.$$

两边取$M \to \infty$并在左边交换和n的极限次序(为什么可以?)便得到定理. 完毕.

它的一个推论是

推论5.1.3. *设$\{f_n\}$是单调的可测函数列, f_1可积. 令$\lim_n f_n = f$ a.e. . 则f的积分存在且*

$$\lim_n \int f_n d\mu = \int f d\mu.$$

证明: 只需证明单调上升的情形, 单调下降时考虑$\{-f_n\}$即可.

f_1可积保证了

$$\int (f_n - f_1) d\mu = \int f_n d\mu - \int f_1 d\mu.$$

于是在上定理中用$f_n - f_1$代替f_n即可. 完毕.

另一个推论是:

推论5.1.4. *(Lebesgue基本定理) 设$\{f_n\}$满足$\sum_n \int |f_n| d\mu < \infty$, 则$f := \sum_n \int f_n$收敛, f可积且*

$$\int |f - \sum_{k=1}^n f_k| d\mu \to 0, \quad n \to \infty.$$

特别地,

$$\int f d\mu = \lim_n \int \sum_{k=1}^n f_k d\mu.$$

证明:

对

$$g_n := \sum_{k=1}^n |f_k|, \quad g := \sum_{k=1}^\infty |f_k|$$

用定理5.1.2得到

$$\int \sum_{k=1}^\infty |f_k| < \infty.$$

因而由命题4.5.2之2),

$$\sum_{k=1}^{\infty}|f_k| < \infty \text{ a.e. }.$$

于是$\sum_{k=1}^{\infty} f_k < \infty$几乎处处收敛. 又由于它不超过可积的$g$, 故也可积, 并且

$$\int |\sum_{k=1}^{\infty} f_k - \sum_{k=n}^{\infty} f_k| d\mu \leq \sum_{k=n+1}^{\infty}\int |f_k| d\mu \to 0 \qquad n\to\infty.$$

完毕.

在继续陈述进一步的结果之前, 我们用现有结果研究一下函数的不定积分. 设f是积分存在的函数, 它的不定积分是指集函数

$$\nu(E) := \int_E f d\mu = \int f 1_E d\mu,\ \forall E \in \mathscr{F}.$$

我们有:

定理5.1.5. *ν是可列可加集函数, 即对任意不交的可测集$\{E_n\}$有*

$$\nu(\cup_n E_n) = \sum_n \nu(E_n)(*).$$

而当f可积时, 此集函数关于μ是绝对连续的, 即$\forall \varepsilon > 0, \exists \delta > 0$ 使$\mu(E) < \delta$时$\nu(E) < \varepsilon$.

证明: 先设f是非负可测函数. 此时在推论5.1.3中以$g := \sum_n f 1_{E_n}$代f, $g_n := \sum_{k=1}^{n} f 1_{E_k}$代$f_n$即得(*). 一般情况分别考虑正、负部即可.

现在假定f可积, 我们来证明绝对连续性. 若否, 则存在$\varepsilon > 0$使对任意自然数n皆存在$E_n \in \mathscr{F}$ 满足$\mu(E_n) < 2^{-n}$ 而

$$\int_{E_n} f d\mu > \varepsilon.$$

令

$$E := \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k.$$

则

$$\mu(E) \leq \lim \sum_{k=n}^{\infty} \mu(E_k) = 0.$$

而

$$\int_E f d\mu = \int f 1_{\cap_{n=1}\infty \cup_{k=n}^{\infty}} E_k d\mu = \lim_n \int f 1_{\cup_{k=n}^{\infty}} E_k d\mu \geq \varepsilon > 0.$$

这是不可能的. 完毕.

好了, 我们回到极限与积分符号的交换问题. 现在的问题是单调性的假设太强, 使其适用范围大受限制, 因而我们要想办法去掉它. 但我们为此要付出什么样的代价呢?

一般说来, 对可积的$\{f_n\}$, f, 总可以写

$$\begin{aligned} & |\int f_n d\mu - \int f d\mu| \\ \leq & \int |f_n - f| d\mu \\ = & \int_{|f_n - f| > \varepsilon} |f_n - f| d\mu + \int_{|f_n - f| \leq \varepsilon} |f_n - f| d\mu \\ \leq & \int_{|f_n - f| > \varepsilon} |f_n| d\mu + \int_{|f_n - f| > \varepsilon} |f| d\mu + \varepsilon\mu(X). \end{aligned} \tag{5.1.15}$$

其中第三项显然可以无条件地任意小——只要ε足够小就行; 第二项由定理5.1.5在n足够大时也可以任意小——如果我们假设f_n几乎处处或依测度收敛到f的话; 要命的是第一项, 我们现有的任何条件都不能保证它任意小——这里问题的关键是尽管积分集合的测度可随n的增大而任意变小, 但被积函数也在随着n变化. 因此, 为保证它任意小, 必须假定f_n的不定积分随着积分集合的变小会一致地变小. 这样就导致了下面的

定义5.1.6. *设I是任意指标集, $\{f_i, i \in I\}$是一族可积函数. 若$\forall \varepsilon > 0$, $\exists \delta > 0$, 使只要$\mu(E) < \delta$, 便有*

$$\sup_{i \in I} \int_E |f_i| d\mu < \varepsilon,$$

则称$\{f_i, i \in I\}$的积分一致绝对连续.

我们现在可以证明

定理5.1.7. *若f可积, $\{f_n\}$的积分一致绝对连续且f_n几乎处处或依测度收敛到f, 则*

$$\lim_n \int |f_n - f| d\mu = 0.$$

证明: $\forall \varepsilon$, δ, 在5.1.15中, 存在N, 当$n > N$时就有

$$\int |f_n - f| d\mu \leq 2\delta + \varepsilon\mu(X).$$

故

$$\limsup_n \int |f_n - f| d\mu \leq 2\delta + \varepsilon\mu(X).$$

由于δ, ε是任意的, 故必有

$$\lim_{n} \int |f_n - f| d\mu = 0.$$

完毕.

推论5.1.8. *(控制收敛定理)若f_n几乎处处或依测度收敛到f且存在可积的F使$f_n \leq F$, $\forall n$, 则上定理的结论成立.*

证明: 由于$\forall E \in \mathscr{F}$,

$$\int_E |f_n| d\mu \leq \int_E |F| d\mu,$$

故f_n一致绝对连续. 完毕.

与绝对连续概念密切相关的是一致可积性. 所谓一致可积性是指:

定义5.1.9. *设T是一指标集,$\{f_t, t \in T\}$是一族可积随机变量. 若*

$$\lim_{c\to\infty} \sup_{t\in T} \int 1_{|f_t|>c} |f_t| d\mu = 0,$$

则称$\{f_t, t \in T\}$一致可积.

"一致可积"其实是一个相当自然的概念. 如果我们将积分与级数对等,那么可积就与绝对收敛——让我们叫做"可和"——对等. 我们回忆一下,为使函数级数的求和号下取极限的运算合法,条件不是别的,就是一致可和.

判断一族随机变量一致可积有下面简单的充分必要条件:

命题5.1.10. *$\{f_t, t \in T\}$一致可积的充要条件是: 存在凸Borel函数$h : \mathbb{R}_+ \to \mathbb{R}_+$满足$\lim_{x\to\infty} \frac{h(x)}{x} = \infty$的,使$\sup_{t\in T} \int h(|f_t|) d\mu < \infty$. 在充分性条件中,函数的凸性可以不要.*

证明: 练习.

习题: 设$\mathscr{H}$为一致可积函数族,定义

$$\mathscr{G} := \overline{\{\sum_{i=1}^n \alpha_i f_i, n \geq 1, \alpha_i \in (-1,1), \sum_{i=1}^n |\alpha_i| = 1\}}.$$

证明$\mathscr{G}$一致可积.

一致可积与一致绝对连续的关系是:

命题5.1.11. *$\{f_t, t \in T\}$一致可积的充要条件是: $\{|f_t|\}$的积分一致有界且一致绝对连续.*

证明:必要性. 设$\{f_t, t\in T\}$一致可积,则$\forall c>0$,

$$\begin{aligned}\int_E |f_t|d\mu &= \int_{E\cap\{|f_t|>c\}} |f_t|d\mu + \int_{E\cap\{|f_t|\le c\}} |f_t|d\mu \\ &\le \int_{\{|f_t|>c\}} |f_t|d\mu + c\mu(E)\end{aligned}$$

所以

$$\limsup_{\mu(E)\to 0}\sup_t \int_E |f_t|d\mu \le \sup_t \int_{\{|f_t|>c\}} |f_t|d\mu$$

令$c\to\infty$,得

$$\lim_{\mu(E)\to 0}\sup_t \int_E |f_t|d\mu = 0.$$

即$\{|f_t|\}$的积分一致绝对连续. 积分的一致有界性由上一命题直接得到.

充分性: $\forall \varepsilon>0$, 取$\delta>0$使

$$\mu(E)<\delta \Longrightarrow \sup_t \int_E |f_t|d\mu < \varepsilon$$

再取$C>0$使

$$\sup_t \mu(|f_t|>C) \le C^{-1}\sup_t \int |f_t|d\mu < \delta$$

则

$$\sup_t \int_{|f_t|>C} |f_t|d\mu < \varepsilon.$$

所以一致可积.

在不满足一致可积的条件时, 我们不可能有积分的收敛. 这时只能转而求其次, 用下面的

引理5.1.12. *(Fatou引理)设g可积, $f_n\ge g$, $\forall n$, 则$\int \liminf_n f_n d\mu \le \liminf_n \int f_n d\mu$.*

证明: 令$g_1=g$, $g_n=\inf_{k\ge n} f_k$, $n\ge 2$, 则

$$\begin{aligned}\int \liminf_n f_n d\mu &= \int \lim_n g_n d\mu \\ &= \lim_n \int g_n d\mu \ \text{由推论5.1.3} \\ &\le \lim_n \inf_{k\ge n} \int f_k d\mu \\ &= \liminf_n \int f_n d\mu.\end{aligned}$$

完毕.

该引理中可积函数g的存在性是必须的. 反例如下: $X=[0,1]$, $\mu=$Lebesgue测度, $f_n(x)=-n^2$, $x\leq n^{-1}$; 0, $x>n^{-1}$.

由于0是可积函数, 因此Fatou引理对非负函数是恒成立的. 特别我们有

推论5.1.13. *设f_n是可测函数, 则*

$$\int \liminf_n |f_n| d\mu \leq \liminf_n \int |f_n| d\mu.$$

§5.2 σ有限测度空间情形

我们来逐条检查上一小节的结果中哪些在σ有限测度空间情形照样成立, 而哪些则不再成立.

有界收敛定理不再成立请举例说明.

定理5.1.7不再成立(请举例说明), 要换为Vitali定理

定理5.2.1. *(Vitali)设$(X,\mathscr{F},\mu)$是σ有限测度空间,$\{f_n\}$是可积函数列, f是可测函数, $f_n\to f$, $\mu-a.e.$. 则f可积且*

$$\lim_n \int |f_n - f| d\mu = 0$$

的充要条件为

(i) $\forall\varepsilon>0$, $\exists A_e\in\mathscr{F}$, 使得$\mu(A_\varepsilon)<\infty$且

$$\int_{A_\varepsilon^c} |f_n| d\mu < \varepsilon, \quad \forall n;$$

(ii)

$$\lim_{\mu(E)\to 0} \sup_n \int_E |f_n| d\mu = 0.$$

其它结果均成立(请证明,尤其是控制收敛定理).

§5.3 应用到带参数的积分

设I是$\mathbb{R}$的一个开区间, $\forall t\in I$, $f(t,\cdot)$是可积函数. 令

$$u(t) := \int f(t,x) d\mu(x).$$

我们来考察u的连续性与可微性问题.

定理5.3.1. *(连续性)设*$t_0 \in I$

1) 对任意数列$t_n \to t$, $t_n \in I$, $f(t_n, x) \to f(t_0, x)$ *a.e.;*

2) 存在g*可积及*$\varepsilon > 0$, *使*$|t - t_0| < \varepsilon$*时有*

$$|f(t,x)| \le g(x) \ \ a.e..$$

则u*在*t_0*点连续.*

证明: u在t_0连续$\Longleftrightarrow \forall t_n \to t_0$, $u(t_n) \to u(t_0)$. 然后用控制收敛定理. 完毕.

定理5.3.2. *(可微性)设*$t_0 \in I$*并假定*

1) 存在零测集E_1, *使*$\forall x \notin K_1$, $\frac{\partial f}{\partial t}(t,x)$*对任意*$t$*存在, 且作为*$t$*的函数在*$t_0$*连续;*

2) 存在g*可积及零测集*K_2, *使*

$$|\frac{\partial f}{\partial t}(t,x)| \le g(x), \quad \forall (t,x) \in I \times K_2^c.$$

则u*在*t_0*处可微且*

$$u'(t_0) = \int_X \frac{\partial f}{\partial t}(t,x) d\mu(x).$$

证明: 只需证明存对任意$\varepsilon_n \to 0$有

$$\lim_n \frac{u(t_0 + \varepsilon_n) - u(t_0)}{\varepsilon_n} = a,$$

其中

$$a := \int_X \frac{\partial f}{\partial t}(t_0, x) d\mu(x),$$

则

$$\frac{u(t_0 + \varepsilon_n) - u(t_0)}{\varepsilon_n} = \int f_n(x) d\mu(x),$$

其中

$$f_n(x) := \frac{f(t_0 + \varepsilon_n, x) - f(t_0, x)}{\varepsilon_n}.$$

再令$K = K_1 \cup K_2$, 则K仍是零测集. 对$x \notin K$, 由中值定理有:

$$f_n(x) = \frac{\partial f}{\partial t}(t_0 + \theta_n(x), x),$$

这里$|\theta_n(x)| < \varepsilon_n$. 于是由1)有

$$\lim_n f_n(x) = \frac{\partial f}{\partial t}(t_0, x), \ \forall x \notin K.$$

又由2)有

$$|f_n(x)| \le g(x) \ \forall x \in K,$$

所以由控制收敛定理即得结论. 完毕.

§5.4 完备化测度空间上的积分

设$(X,\mathscr{F},\mu)$为测度空间, $(X,\tilde{\mathscr{F}},\tilde{\mu})$为其完备化. 我们已经知道, $\mathscr{F}$中的元素与$\tilde{\mathscr{F}}$中的元素并没有太大的差别——一个零测集的子集而已. 我们现在说明对积分也是如此. 为此先证明:

命题5.4.1. *设$\tilde{f}\in\tilde{\mathscr{F}}$, 则$\exists f\in\mathscr{F}$使$f=\tilde{f}$, $\tilde{\mu}$ a.e..*

证明: 记

$$\mathcal{L}:=\{\tilde{f}\in\tilde{\mathscr{F}}:\exists f\in\mathscr{F}\text{ 使}f=\tilde{f},\tilde{\mu}a.e..\}$$

则$1_E\in\mathcal{L}$, $\forall E\in\tilde{\mathscr{F}}$, 且$\mathcal{L}$是线性空间, 对单调极限封闭, 因此由定理2.6.11知$\mathcal{L}$等于$\tilde{\mathscr{F}}$可测函数全体. 完毕.

保持这个定理的记号, 用同样的方法可以证明:

定理5.4.2. *$\tilde{f}$关于$\tilde{\mu}$的可积(积分存在)$\Longrightarrow f$关于μ的可积(积分存在).*

6 变量代换

我们都知道变量代换在计算Riemann积分时的重要作用. 同样的道理, 在Lebesgue积分中焉能够没有变量代换公式! 本节的目的就是为了得到这样一个公式. 可是我们必须注意到, 由于我们现在是在抽象空间中讨论积分, 我们没有了坐标, 没有了连续性, 没有了可微性, 更谈不上Jacobi行列式了——我们什么都没有了, 我们该怎么办呢?

§6.1 映射的逆像

既然什么都没有了, 我们干脆白手起家. 正如代替连续函数的是可测函数一样, 代替连续映射的将是可测映射. 但是且慢, 我们要从有关映射的一般的概念说起.

设X,Y是两空间, ϕ为X到Y的映射.

定义6.1.1. *$\forall E\subset Y$, E在ϕ下的逆像为*

$$\phi^{-1}(E):=\{x\in X:\phi(x)\in E\}.$$

注意这里的ϕ^{-1}并不是我们过去所熟悉的逆映射——后者只对内射才有意义. 实际上它根本就不是点到点的映射, 而是集合到集合的映射. 无论对什么样的ϕ, ϕ^{-1}总是可以定义的.

ϕ^{-1}的一个简单而有用的特点是它保持所有的集合运算关系不变, 也就是说, 下面的命题成立. 其证明是直接的, 故略去.

命题6.1.2. 设I是任意指标集, $E, E_1, E_2, E_i (i \in I) \subset Y$. 则

i) $\phi^{-1}(Y) = X, \phi^{-1}(\emptyset) = \emptyset$;

ii) $E_1 \subset E_2 \Longrightarrow \phi^{-1}(E_2) \subset \phi^{-1}(E_2)$;

iii) $\phi^{-1}(\cap_{i\in I} E_i) = \cap_{i\in I}\phi^{-1}(E_i)$;

iv) $\phi^{-1}(\cup_{i\in I} E_i) = \cup_{i\in I}\phi^{-1}(E_i)$;

v) $\phi^{-1}(E_2 - E_1) = \phi^{-1}(E_2) - \phi^{-1}(E_1)$.

我们再引进记号: 设$\mathcal{E}$是一Y的子集类, 记

$$\phi^{-1}(\mathcal{E}) := \{\phi^{(-1)}E : E \in \mathcal{E}\}.$$

由上一命题知道, 环、代数、σ代数的逆像依然是环、代数、σ代数. 此外, 我们还容易证明

命题6.1.3. 设$\mathcal{E}$是一集类, 则

$$\phi^{-1}(\mathcal{A}(\mathcal{E})) = \mathcal{A}(\phi^{-1}(\mathcal{E}));\quad \phi^{-1}(\sigma(\mathcal{E})) = \sigma(\phi^{-1}(\mathcal{E})).$$

§6.2 可测映射

现在设X, Y带有可测结构, 即$(X, \mathscr{F})$与$(Y, \mathscr{G})$均为可测空间.

定义6.2.1. $\phi : X \mapsto Y$若满足

$$\phi^{-1}(\mathscr{G}) \subset \mathscr{F},$$

则称为$\mathscr{F}/\mathscr{G}$可测, 简记为$\phi \in \mathscr{F}/\mathscr{G}$.

我们来证明可测映射的复合是可测的, 即

定理6.2.2. 设$(X_i, \mathscr{F}_i), i = 1, 2, 3$为可测空间, $\phi_i \in \mathscr{F}_i/\mathscr{F}_{i+1}, i = 1, 2$, 则$\phi_2 \circ \phi_1 \in \mathscr{F}_1/\mathscr{F}_3$.

证明:

$$(\phi_2 \circ \phi_1)^{-1}(\mathscr{F}_i) = \phi_1^{-1}(\phi_2^{-1}(\mathscr{F}_3)) \subset \phi_1^{-1}(\mathscr{F}_2) \subset \mathscr{F}_1.$$

完毕.

下面考虑另一类问题. 设X_2、X_3带有可测结构$\mathscr{F}_2$、$\mathscr{F}_3$, 而X上是空白的. 现在我们考虑映射$\phi_1 : X_1 \mapsto X_2$所产生的X_1上的σ代数

$$\sigma(\phi_1) := \phi_1^{-1}(\mathscr{F}_2).$$

这时按上一定理, 对任意$\phi_2 \in \mathscr{F}_2/\mathscr{F}_3$, $\phi_2 \circ \phi_1 \in \sigma(\phi_1)/\mathscr{F}_3$. 有趣的是, 在某些情况下反过来也成立.

定理6.2.3. *设$(Y,\mathscr{G})$为可测空间, ϕ为映射:$X \mapsto Y$. 则对$(X,\phi^{-1}(\mathscr{G}))$上的任意可测函数$f$, 恒存在$(Y,\mathscr{G})$上的可测函数$g$使$f = g \circ \phi$.*

证明: 记这样的函数类为$\mathcal{H}$. 显然H为线性空间. 现在证明它是单调类. 设$f_n \uparrow f$, $f_n \in \mathcal{H}$, 相应的Y上的函数记为g_n. 自然这时g_n未必是单调的, 但若令

$$h_n := \vee_{k=1}^n g_n,$$

则h_n是单调的且$\lim_n f_n = (\lim_n h_n) \circ \phi$, 因此$f \in \mathcal{H}$.

又对$f = 1_{\phi^{-1}(E)}$, $E \in \mathscr{G}$, 取$g = 1_E$知$f \in \mathcal{H}$. 因此由单调类定理即得结论.

§6.3 变量代换公式

现在设$(X,\mathscr{F},\mu)$是σ有限测度空间, 即有$X_n \in \mathscr{F}, n \geq 1$使$\cup_{n\geq 1} X_n = X$且$\mu(X_n) < \infty$, $\forall n$; $(Y,\mathscr{G})$是可测空间; $\phi : X \mapsto Y$, $\phi \in \mathscr{F}/\mathscr{G}$. 在$\mathscr{G}$上定义集函数$\mu\phi^{-1}$如下:

$$\mu\phi^{-1}(E) := \mu(\phi^{-1}(E)).$$

容易验证$\mu\phi^{-1}$是可列可加的. 但$\mu\phi^{-1}$未必是σ有限的. 不过, 必要时我们将

$$Y_n := \phi(X_n),\ n \geq 1,$$

添加到$\mathscr{G}$中, 则在$\sigma(\mathscr{G}, Y_n, n \geq 1)$上$\mu\phi^{-1}$是$\sigma$有限的, 因为此时

$$\mu\phi^{-1}(Y_n) = \mu(X_n) < \infty.$$

这样我们干脆一开始就直接假定$\mu\phi^{-1}$是σ有限的. 于是, 关于μ与$\mu\phi^{-1}$的积分都是可以定义的.

本小节的主要结果是:

定理6.3.1. *对$(Y,\mathscr{G})$上的任意可测函数f有*

$$\int_X f(\phi) d\mu = \int_Y f d\mu\phi^{-1}.$$

上式的意义是: 一边存在时, 另一边也存在且两边相等.

证明: 首先不失一般性可以假定μ是有限的, 因为否则可以限制在每个X_n上考虑.

其次, 由积分的定义, 只需对非负的f证明.

再次, 还是由积分的定义, 只需对非负的简单函数f证明.

最后, 由积分的线性性, 只要对示性函数证明——而这是显然的. 完毕.

7 乘积空间

§7.1 集合的乘积

不管你是否意识到了, 你实际上已经多次碰到乘积空间: 数学分析里的$\mathbb{R}^n$是n个$\mathbb{R}^1$的乘机; 线性代数里的n维线性空间是n个一维线性空间的乘积; 还有乘积拓扑空间, 等等. 我们先从纯粹集合论的观点固定以后要用的一些基本概念.

定义7.1.1. 设E_1, E_2为两集合, 定义其乘积为集合

$$E_1 \times E_2 := \{(x_1, x_2) : x_1 \in E_1, x_2 \in E_2\};$$

类似地, 设$E_1, \cdots, E_n$为n个集合, 定义其乘积为集合

$$\Pi_{i=1}^n E_i := \{(x_1, \cdots, x_n) : x_1 \in E_1, x_n \in E_n\}.$$

在上面的定义中, 严格地讲, 乘积是与次序有关的: 例如$E_1 \times E_2$和$E_2 \times E_1$是不一样的——除非$E_1 = E_2$. 但是, 它们之间显然按自然方式即(x_1, x_2)对应于(x_2, x_1)的方式是一一对应的, 因此我们以后就认为它们相等. 同样地, 在例如$(E_1 \times E_2) \times E_3$与$E_1 \times E_2 \times E_3$之间也存在自然的一一对应, 因此也认为它们相等. 这样, 我们就可以给出任意多个集合的乘积的定义.

定义7.1.2. 设I是任意指标集, 而对每一$i \in I$, E_i是一集合, 则定义其乘积为

$$\begin{aligned} \Pi_{i\in I} : &= \{\{x_i, i \in I\} : x_i \in E_i \ \forall i\} \\ &= \{x : x : I \mapsto \cup_{i\in I} E_i, \ x(i) \in E_i\}. \end{aligned}$$

当诸E_i就是全空间X_i时, 所得到的乘积集合便很自然称为乘积空间. 我们现在从最简单的二维乘积空间开始介绍将要用到的一些概念. 注意这里的维数纯粹是指作成乘积空间时的因子空间的数目, 与分析中说的维数没有任何关系——比如说, 当我们要将$\mathbb{R}^3$ 看成$\mathbb{R}^1 \times \mathbb{R}^2$时, 它就是二维的.

由于一般有限维乘积空间与二维乘积空间相比, 除了记号复杂一些外, 并没有什么本质的不同, 因此我们将着重研究二维情形.

这样, 我们设有两空间X, Y. 我们以后把$X \times Y$的形如

$$A \times B, A \subset X, B \subset Y$$

的子集称为以A, B为边的矩形.

下面是关于矩形的一些简单事实.

定理7.1.3. *i) 当且仅当一条边是空集时, 矩形是空集;*

ii) 设$E_1 = A_1 \times B_1$, $E_2 = A_2 \times B_2$, 则当且仅当$A_1 \subset A_2$, $B_1 \subset B_2$时, $E_1 \subset E_2$;

iii) 设$E_1 = A_1 \times B_1$, $E_2 = A_2 \times B_2$, 则当且仅当$A_1 = A_2$, $B_1 = B_2$时, $E_1 = E_2$;

证明:

i) 设$A \times B \neq \emptyset$, 则$\exists (x, y) \in A \times B$, 所以$x \in A$, $y \in B$, 因而A、B均不为空集. 反之, 设A、B均不为空集, 则$\exists x \in A$, $y \in B$, 所以$(x, y) \in A \times B$, 即$A \times B$不为空集.

ii) "当"部分显然; "仅当"部分证明如下: 设$E_1 \subset E_2$, 则$\forall x \in A_1$, $y \in B_1$, $(x, y) \in E_1 \subset E_2$, 因而$x \in A_2$, $y \in B_2$.

iii) 直接由ii). 完毕.

§7.2 乘积可测结构

设给定了两个可测空间$(X, \mathscr{F})$与$(Y, \mathscr{G})$, 我们的目的是要在乘积空间$(X \times Y)$上构造一个σ代数, 它与因子空间上已有的可测结构自然相连.

我们将形如$A \times B$, 其中$A \in \mathscr{F}$, $B \in \mathscr{G}$, 的集合称为可测矩形. 它们的全体显然构成一π类.

定义7.2.1. *定义*

$$\mathscr{F} \times \mathscr{G} := \sigma(\text{ 全体可测矩形}).$$

称$(X \times Y, \mathscr{F} \times \mathscr{G})$为乘积可测空间.

注意, 由于可测矩形全体成一π类, 故它产生的λ类也是$\mathscr{F} \times \mathscr{G}$.

下面我们要涉及一个重要概念, 这就是截口.

定义7.2.2. *设$E \subset X \times Y$, 对任意$x \in X$, 定义*

$$E_x := \{y \in Y : (x, y) \in E\};$$

对任意$y \in Y$, 定义

$$E^y := \{x \in X : (x, y) \in E\}.$$

他们分别称为E在x及y处的截口.

容易直接验证, 截口满足如下关系:

$$(E - F)_x = E_x - F_x, \tag{7.2.16}$$

$$(\cup_{i\in I}E_i)_x=\cup(E_i)_x,\quad (\cap_{i\in I}E_i)_x=\cap(E_i)_x, \tag{7.2.17}$$

以及关于$y-$截口的同样的关系式.

显然, 矩形的截口总是它的某条边, 因此可测矩形的截口是因子空间的可测集. 这一结论可推广到一般的可测集.

定理7.2.3. *设$E\in\mathscr{F}\times\mathscr{G}$, 则$\forall x\in X$, $y\in Y$, 有$E_x\in\mathscr{G}$, $E^y\in\mathscr{F}$.*

证明: 令

$$\mathcal{A}:=\{E\in\mathscr{F}\times\mathscr{G}:E_x\in\mathscr{G},E^y\in\mathscr{F}\}.$$

则$\mathcal{A}$包含了所有可测矩形这一π类, 再由7.2.16与7.2.17知$\mathcal{A}$是λ类, 因此由定理2.6.10$\mathcal{A}=\mathscr{F}\times\mathscr{G}$. 完毕.

对于$X\times Y$上的函数f, 我们可类似地定义其截口为

$$f_x(\cdot):=f(x,\cdot),\quad f^y(\cdot):=f(\cdot,y).$$

我们有:

定理7.2.4. *设$f\in\mathscr{F}\times\mathscr{G}$, 则$\forall x\in X$, $y\in Y$, 有$f_x\in\mathscr{G}$, $f^y\in\mathscr{F}$.*

证明: 对任意Borel集B有

$$f_x^{-1}(B)=(f^{-1}(B))_x\in\mathscr{G}.$$

所以$f_x\in\mathscr{G}$. 同理$f^y\in\mathscr{F}$. 完毕.

习题:

1. 证明: 可测的矩形是可测矩形.

§7.3　乘积测度

现设$(X,\mathscr{F},\mu)$与$(Y,\mathscr{G},\nu)$均为有限测度空间. 我们来证明

定理7.3.1. *在$(X\times Y,\mathscr{F}\times\mathscr{G})$上存在唯一一个有限测度, 记为$\mu\times\nu$, 使$\forall A\in\mathscr{F}$, $B\in\mathscr{G}$有*

$$\mu\times\nu(A\times B)=\mu(A)\times(B).$$

证明: 先证明唯一性. 设l与l'均是$\mathscr{F}\times\mathscr{G}$上的有限测度且在可测矩形上都相等. 令

$$\mathcal{L}:=\{E\in\mathscr{F}\times\mathscr{G}:l(E)=l'(E)\}.$$

易证$\mathcal{L}$为λ类. 于是$\mathcal{L}=\mathscr{F}\times\mathscr{G}$.

次证存在性. 为此要证明:

引理7.3.2. $\forall E \in \mathscr{F} \times \mathscr{G}$, 有$\nu(E_x) \in \mathscr{F}$, $\mu(E^y) \in \mathscr{G}$且

$$\int_X \nu(E_x)\mu(dx) = \int \mu(E^y)\nu(dy).$$

证明: 令

$$\mathcal{L} := \{\forall E \in \mathscr{F} \times \mathscr{G} : \nu(E_x) \in \mathscr{F}, \mu(E^y) \in \mathscr{G}, \int_X \nu(E_x)\mu(dx) = \int \mu(E^y)\nu(dy)\}.$$

则$\mathcal{L}$包含了可测矩形全体且为λ类, 因此$\mathcal{L} = \mathscr{F} \times \mathscr{G}$.

现在回到定理的证明. $\forall E \in \mathscr{F} \times \mathscr{G}$令

$$\mu \times \nu(E) = \int_X \nu(E_x)\mu(dx) = \int \mu(E^y)\nu(dy).$$

我们要证明$\mu \times \nu$是有限测度. 这只要证证可列可加性, 因为其它几条性质都是显然的.

设诸E_i不交, 我们有

$$\begin{aligned}
\mu \times \nu(\cup_{i=1}^{\infty} E_i) &= \int_X \nu((\cup_{i=1}^{\infty} E_i)_x)\mu(dx) \\
&= \int_X \sum_{i=1}^{\infty} \nu((E_i)_x)\mu(dx) \text{ 测度的可列可加性} \\
&= \sum_{i=1}^{\infty} \int_X \nu((E_i)_x)\mu(dx) \text{ Lebesgue定理} \\
&= \sum_{i=1}^{\infty} \mu \times \nu(E_i).
\end{aligned}$$

可列可加性得证. 完毕.

习题: 证明在μ及ν为σ有限时, 结论仍然成立.

§7.4 Fubini定理

所谓Fubini定理即重积分化为累次积分的定理. 在数学分析中我们曾经首次碰到过、无穷次使用过关于Riemann积分的这样的定理. 现在我们要在我们刚刚建立的积分理论框架中证明这个定理.

为了时刻提醒$(X \times Y, \mathscr{F} \times \mathscr{G}, \mu \times \nu)$是乘积测度空间, 对$X \times Y$上的函数$f \in \mathscr{F} \times \mathscr{G}$我们用重积分符号

$$\iint_{X \times Y} f d(\mu \times \nu)$$

表示其的积分.

定理7.4.1. 设$f(x,y)$是$X\times Y$上的$\mathscr{F}\times\mathscr{G}$-非负可测函数, $\forall x\in X,\ y\in Y$, 令

$$g(x):=\int_Y f(x,y)\nu(dy),\quad h(y):=\int_X f(x,y)\mu(dx).$$

则$g\in\mathscr{F}$, $\in\mathscr{G}$且

$$\iint_{X\times Y} fd(\mu\times\nu)=\int_X g(x)\mu(dx)=\int_Y h(y)\nu(dy).$$

证明: 以$\mathcal{L}$表示使上式成立的有界$\mathscr{F}\times\mathscr{G}$-可测函数全体, 则$\mathcal{L}$包含了所有可测矩形的示性函数, 且$\mathcal{L}$满足定理2.6.11 的全部条件, 因此$\mathcal{L}$就是有界$\mathscr{F}\times\mathscr{G}$-可测函数全体.

对任意非负$\mathscr{F}\times\mathscr{G}$-可测函数, 可取有界$\mathscr{F}\times\mathscr{G}$-可测函数列单调上升地逼近, 于是由单调收敛定理知等式成立. 完毕.

在未必非负的情形, 有

推论7.4.2. 设f的积分存在, 则仍有

$$\iint_{X\times Y} fd(\mu\times\nu)=\int_X g(x)\mu(dx)=\int_Y h(y)\nu(dy).$$

证明: 直接由

$$\iint_{X\times Y} fd(\mu\times\nu)=\iint_{X\times Y} f^+d(\mu\times\nu)-\iint_{X\times Y} f^-d(\mu\times\nu)$$

及上一定理推出. 完毕.

最后我们指出, 用归纳法可将上面的所有结论推广到任意有限维乘积空间的情形, 我们就不一一叙述了.

习题:

1. 证明在μ及ν为σ有限时, 结论仍然成立.

2. 设$(X,\mathscr{F},\mu)$与$(Y,\mathscr{G},\nu)$均为完备的σ有限测度空间, $f\in\widetilde{(\mathscr{F}\times\mathscr{G})}^{\mu\times\nu}$为$\mu\times\nu$-积分存在, 证明: 对$\mu$-几乎所有的$x\in X$, $y\in Y$, 有$f_x\in\mathscr{G}$, $f^y\in\mathscr{F}$; $\int_Y f_xd\nu$, $\int_X f^yd\mu$存在且

$$\iint_{X\times Y} fd\widetilde{(\mu\times\nu)}=\int_Y\mu(dy)\int_X f(x,y)\nu(dx)=\int_Y\nu(dy)\int_X f(x,y)\mu(dx).$$

3.设$(\Omega,\mathscr{F},P)$是概率空间, ξ是定义在其上的非负随机变量. 证明

$$E[\xi^p]=p\int_0^\infty t^{p-1}P(\xi>t)dt,\quad\forall p>0.$$

8 无限维乘积空间

上一节的结果给我们提供了构造任意有限个具有给定分布的独立随机变量的工具. 但若要构造一列具有给定分布的独立随机变量序列, 或构造具有给定有限维分布的随机过程, 它就不够用了. 这时需要无限维乘积空间.

我们将只考虑概率空间, 这就相当要求全空间的测度为1. 这样要求的原因有二: 一是我们的目的直接是概率论, 因此只需要考虑概率空间; 其次也巧, 从本质上说, 不是概率空间还真无法定义无穷乘积——这从我们下面的构造过程可清楚地看出.

§8.1 可列无限维乘积空间

我们设$(X_i, \mathscr{F}_i, P_i)$, $i = 1, 2, \cdots$是一列可测空间.

定义8.1.1. *设$E \subset \Pi_{i=1}^{\infty} X_i$且存在$n$, $A \in \Pi_{i=1}^{n} \mathscr{F}_i$使$E = A \times \Pi_{i=n+1}^{\infty} X_i$, 则$E$称为(以$A$ 为底的)可测柱集. 由可测柱集全体产生的σ代数称为$\Pi_{i=1}^{\infty} X_i$上的乘积σ代数, 记为$\Pi_{i=1}^{\infty} \mathscr{F}_i$.*

现在设P_i是$(X_i, \mathscr{F}_i)$上的概率测度, 在$\Pi_{i=1}^{\infty} \mathscr{F}_i$上定义集函数$P$如下:

$$E = A \times \Pi_{i=n+1}^{\infty} X_i, \ A \in \Pi_{i=1}^{n} \mathscr{F}_i$$

时

$$P(E) := (\Pi_{i=1}^{n} P_i)(A).$$

这一定义显然是无歧义的: 若$E = A \times \Pi_{i=n+1}^{\infty} X_i = B \times \Pi_{i=m+1}^{\infty} X_i$. 则$n = m$时必有$A = B$; $n > m$时必有$A = B \times \Pi_{i=m+1}^{n} X_i$. 无论如何都有$(\Pi_{i=1}^{n} P_i)(A) = (\Pi_{i=1}^{m} P_i)(B)$.

可测柱集全体显然构成一代数, 我们记它为$\mathcal{A}$. 我们来证明:

定理8.1.2. *P是$\mathcal{A}$上测度.*

证明: 除开可列可加性以外, 其它都是显然的. 而为证可列可加性, 只需证P在$\emptyset$处是上连续的.

设$E_n \in \mathcal{A}$, $E_n \downarrow \emptyset$, 我们要证明$P(E_n) \downarrow 0$.

令

$$Q_n := \Pi_{i=1}^{n} P_i.$$

必要时重复若干项, 可不妨假定$E_n = F_n \times (\Pi_{i=n+1}^{\infty} X_i)$, 其中$F_n \in \Pi_{i=1}^{n} \mathscr{F}_i$.

设有$\varepsilon > 0$使$P(E_n) \geq \varepsilon$, $\forall n$. $\forall x_1 \in X_1$, 令

$$F_n(x_1) := \{(x_2, x_3, \cdots, x_n) \in \Pi_{i=2}^{n} X_i : (x_1, x_2, \cdots, x_n) \in F_n\}.$$

由于E_n单调下降, 所以$x_1 \notin F_1$时$F_n(x_1) = \emptyset,\ n \geq 2$. 于是由Fubini定理有

$$\varepsilon < P(E_n) = Q_n(F_n) = \int_{X_1} (\Pi_{i=2}^n P_i)(F_n(x_1)) P_1(dx_1) = \int_{F_1} (\Pi_{i=2}^n P_i)(F_n(x_1)) P_1(dx_1)$$

令

$$G_n := \{x_1 \in F_1 : (\Pi_{i=2}^n P_i)(F_n(x_1)) > \varepsilon/2\}.$$

则

$$\varepsilon \leq P_1(G_n) + \frac{\varepsilon}{2}$$

故

$$P_1(G_n) > \varepsilon/2.$$

显然G_n是单调下降的, 因而由P_1的上连续性知存在$x_1^0 \in F_1$使$x_1^0 \in \cap_{n=1}^\infty G_n$, 即

$$(\Pi_{i=2}^n P_i)(F_n(x_1^0)) > \varepsilon/2.$$

现在, 令$E_n(x_1^0) := F_n(x_1^0) \times \Pi_{i \geq n} X_i$并以$(\varepsilon, \Pi_{n\geq 2} X_n, \Pi_{n\geq 2} \mathscr{F}_n, \Pi_{n\geq 2} P_n, E_n(x_1^0))$代$(\varepsilon/2, \Pi_{n\geq 1} X_n, \Pi_{n\geq 1} \mathscr{F}_n, \Pi_{n\geq 1} P_n, E_n)$而重复以上推理, 可知有$x_2^0 \in F_2(x_1^0)$——因而$(x_1^0, x_2^0) \in F_2$——使

$$(\Pi_{i=3}^n P_i)(F_n(x_1^0, x_2^0)) > \varepsilon/4,\ n \geq 3$$

这里

$$F_n(x_1, x_2) := \{(x_3, \cdots, x_n) \in \Pi_{i=3}^n X_i : (x_1, x_2, \cdots, x_n) \in F_n\}.$$

依次下去, 我们便可得到一列$\{x_n^0, n \geq 1\}$使

$$(x_1^0, \cdots, x_n^0) \in F_n.$$

于是自然有

$$(x_1^0, x_2^0, \cdots) \in E_n, \quad \forall n.$$

故

$$\cap_{n=1}^\infty E_n \neq \emptyset.$$

完毕.

在无穷维可测矩形上, 乘积测度确实是一个"乘起来"的测度:

命题8.1.3. 设$E_n \subset X_n,\ E_n \in \mathscr{F}_n$, 则

$$(\Pi_{n=1}^\infty P_n)(\Pi_{n=1}^\infty E_n) = \Pi_{n=1}^\infty P(E_n).$$

证明: 令

$$F_n = (\Pi_{i=1}^n E_i) \times \Pi_{i=n+1}^\infty X_i,$$

则$F_n \downarrow \Pi_{n=1}^\infty E_n$. 但

$$(\Pi_{n=1}^\infty P_n)(F_n) = (\Pi_{i=1}^n P_i)(\Pi_{i=1}^n E_i),$$

故

$$(\Pi_{n=1}^\infty P_n)(\Pi_{n=1}^\infty E_n) = \lim_{n\to\infty} \Pi_{i=1}^n P(E_i) = \Pi_{n=1}^\infty P(E_n).$$

完毕.

§8.2 可列乘积可测空间上的非乘积测度

设P是$(\mathbb{R}^\infty, \mathscr{B}^\infty)$上的概率测度. 定义$\mathbb{R}^\infty$到$\mathbb{R}^n$的投影映射

$$\pi_n((x_1, x_2, \cdots)) = (x_1, x_2, \cdots, x_k).$$

则$P_n := \pi_n^{-1}$是$\mathbb{R}^n$上的测度,且满足下列相容性条件

$$P_n = P_m \pi_{m,n}^{-1}$$

其中$m > n$, $\pi_{m,n}$是从$\mathbb{R}^m$到$\mathbb{R}^n$的自然投影:取前n个坐标.

下面提出反问题: 若在每个$\mathbb{R}^n$上都有概率测度P_n且它们满足相容性条件,是否在$\mathbb{R}^\infty$上存在概率测度P使$P_n = P\pi_n^{-1}$. 这个问题及下节在不可列乘积空间上的同样问题涉及满足给定有限维分布的随机过程的存在性问题,是由—-再一次是由—-Kolmogorov解决的.

定理8.2.1. 设$P_1, P_2, \cdots$分别是$(\mathbb{R}^1, \mathscr{B}^1), (\mathbb{R}^2, \mathscr{B}^2), \cdots$上的概率测度并满足上述相容性条件, 则在$(\mathbb{R}^\infty, \mathscr{B}^\infty)$上存在唯一一个概率测度$P$满足

$$P\pi_n^{-1} = P_n.$$

证明: 沿用上面的记号, 对$B \in \mathscr{B}^n$, 令

$$P(\pi_n^{-1}(B)) = P_n(B).$$

由相容性条件, P是$\mathscr{A}$上有确切定义的集函数. 显然P是有限可加的.

下面证明它也是可列可加的. 为此只需证明它在$\emptyset$处是上连续的. 设$B_n' \in \mathscr{A}$, $B_n' \downarrow \emptyset$. 必要时添加或减少一些元素,可设存在$B_n \in \mathscr{B}^n$使

$$B_n' = \pi_n^{-1}(B_n).$$

设$\varepsilon > 0$使

$$P(B_n') > \varepsilon, \quad \forall n.$$

取紧集$K_n \subset B_n$使

$$P_n(B_n - K_n) \leq \varepsilon 2^{-n}.$$

令

$$K_n' = \pi_n^{-1}(K_n)$$

则

$$B_n' - K_n' = \pi_n^{-1}(B_n - K_n)$$

所以

$$P(B_n' - K_n') = P_n(B_n - K_n) \leq \varepsilon 2^{-n}.$$

令

$$D_n := \cap_{k=1}^n K_k.$$

$$D_n' = \pi_n(D_n)$$

由于$B_n' \downarrow$, 所以

$$B_n' - D_n' = \cup_{k=1}^n (B_n' - K_k') \subset \cup_{k=1}^n (B_k' - K_k')$$

从而

$$P(B_n' - D_n') \leq \sum_{k=1}^n P(B_n' - D_n') \leq \frac{\varepsilon}{2}$$

再由假定$P(B_n') > \varepsilon$得

$$P(D_n') > \varepsilon.$$

因而对任意n, $D_n' \neq \emptyset$. 故可取$x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \cdots) \in D_n'$. 因为$\{x_1^{(n)}\} \subset K_1$而$K_1$紧, 故可抽取子列$n_1$及$x_1^0 \in K_1$使

$$x_1^{(n_1)} \to x_1^0.$$

再从$\{n_1\}$中取子列$\{(n_2)\}$使

$$(x_1^{(n_2)}, x_2^{(n_2)}) \to (x_1^0, x_2^0) \in K_2.$$

依次下去, 对任意k有第k级子列$\{(n_k)\}$使

$$(x_1^{(n_k)}, x_2^{(n_k)}, \cdots, x_k^{(n_k)}) \to (x_1^0, x_2^0, \cdots, x_k^0) \in K_k.$$

用对角线法, 最终可抽取子列$\{n_k\}$使

$$(x_1^{(m_k)}, x_2^{(m_k)}, \cdots) \to (x_1^0, x_2^0, \cdots) \in K_k.$$

显然$(x_1^0, x_2^0, \cdots) \in \cap_n D_n'$. 但$\cap_n D_n' \subset \cap_n B_n'$, 故$\cap B_n' \neq \emptyset$. 证毕.

比较该定理和上一定理,可以看到这里作为乘积因子我们是取了特殊的可测空间$(\mathbb{R}, \mathscr{B})$. 不过我们注意到我们只是用了$(\mathbb{R}^n, \mathscr{B}^n)$上的概率测度是所谓“胎紧的”这一性质, 而该性质对所有Polish空间上的概率测度均成立, 因此定理可以推广到每个乘积因子是Polish空间(带Borel代数)的情形,证明不需要做任何改变.

然而我们不能不要任何拓扑假设. 这方面的反例可见SHIRYAEV p.165.

§8.3 任意无限维乘积空间

现在设T为一不可数无穷集合——应用中最常见的是取T为$\mathbb{R}$的某个区间, $t \in T$表示时间. 设对任意$t \in T$, $(X_t, \mathscr{F}_t)$为可测空间.

首先, 类似于定义8.1.1, 我们给出

定义8.3.1. *设$E \subset \Pi_{t\in T} X_t$且存在$T$的有限子集$S$, $A \in \Pi_{t\in S}\mathscr{F}_t$使$E = A \times \Pi_{t\in S^c} X_t$, 则$E$称为(以$A$为底的)可测柱集. 由可测柱集全体产生的$\sigma$代数称为$\Pi_{t\in T} X_t$上的乘积$\sigma$代数, 记为$\Pi_{t\in T}\mathscr{F}_t$.*

下面是一个关键的结果, 利用它我们可以看出, 无论是什么样的"无穷维"乘积, 在某种意义下都可以转化为"可列无穷维"乘积.

定理8.3.2.

$$\Pi_{t\in T}\mathscr{F}_t = \bigcup_{S\subset T, S \text{ 可数}} \Pi_{t\in S}\mathscr{F}_t.$$

证明: 上式右边显然包含于左边; 又右边为σ代数, 且包含了所有可测柱集, 所以它包含了左边. 完毕.

现在设每一$(X_t, \mathscr{F}_t)$上均有一概率测度μ_t, 在$\Pi_{t\in T}\mathscr{F}_t$上定义

$$(\Pi_{t\in T}\mu_t)(E) := \Pi_{t\in S}\mu_t(A), \quad 若 E = A \times \Pi_{t\in S^c} X_t,\ A \in \Pi_{t\in S}\mathscr{F}_t, \quad S \text{ 可数}.$$

我们要说明该定义是无歧异的.

设同时有

$$E = A_1 \times \Pi_{t\in S_1^c} X_t = A_2 \times \Pi_{t\in S_2^c} X_t,$$

其中S_1与S_2均为可数集且$A_1 \in \Pi_{t\in S_1}\mathscr{F}_t$、$A_2 \in \Pi_{t\in S_2}\mathscr{F}_t$. 令

$$S := S_1 \cap S_2,$$

$A := E$ 在$\Pi_{t\in S} X_t$ 上的投影 $= A_1$ 在$\Pi_{t\in S} X_t$ 上的投影 $= A_2$ 在$\Pi_{t\in S} X_t$ 上的投影. 则

$$\Pi_{t\in S_1}\mu_t(A_1) = \Pi_{t\in S}\mu_t(A)\Pi_{t\in S_1 - S}\mu_t(X_t) = \Pi_{t\in S}\mu_t(A)\Pi_{t\in S_2 - S}\mu_t(X_t) = \Pi_{t\in S_2}\mu_t(A_2)$$

因此定义的确是无歧异的.

定理8.3.3. *以上定义的μ是概率测度.*

证明: 除可列可加性以外, 都是显然的. 而可列可加性可如下证明: 设$E_n \in \Pi_{t\in T}\mathscr{F}_t$, 则存在$S_n \subset T$, S_n 可数, 使$E_n \in \Pi_{t\in S_n}\mathscr{F}_t$. 因而对所有的$n$, $E_n \in \Pi_{t\in S}\mathscr{F}_t$, 其中$S = \cup_{n=1}^{\infty} S_n$. 这样, 利用$\mu$在$\Pi_{t\in S}\mathscr{F}_t$上的可列可加性即得. 完毕.

§8.4 任意无限维乘积可测空间上的非乘积测度

对任意$t \in T$, 令

$$\mathbb{R}_t = \mathbb{R}^1, \quad \mathscr{B}_t = \mathscr{B}^1$$

与上面的推理类似,从定理8.2.1出发,可以得到:

定理8.4.1. *对任意有限子集$\tau = \{t_1, \cdots, t_n\} \subset T$, 令*

$$(\mathbb{R}^\tau, \mathscr{B}^\tau) = (\mathbb{R}_{t_1}, \mathscr{B}_{t_1}) \times \cdots \times (\mathbb{R}_{t_n}, \mathscr{B}_{t_n})$$

设P_τ为$\mathscr{B}^\tau$上的测度, 满足下列相容性条件:

$$\tau_1 \subset \tau_2 \Longrightarrow P_{\tau_1} = P_{\tau_2} \pi^{-1}_{\tau_2, \tau_1}$$

其中π_{τ_2,τ_1}是$\mathbb{R}^{\tau_2}$到$\mathbb{R}^{\tau_1}$的自然投影.则在$(\mathbb{R}^T, cB^T)$上存在唯一一个概率测度P^T使

$$P_\tau = P\pi_\tau^{-1}.$$

9 在概率论上的应用

我们将把现有的知识应用到概率论上. 具体地说, 给定一个分布函数或一个分布函数列, 我们要构造一个概率空间及其上的一个随机变量或独立随机变量序列, 服从所给定的分布.

10 L^p空间

我们固定一个σ有限测度空间$(X, \mathscr{F}, \mu)$.

§10.1 定义及基本不等式

定义10.1.1. *设$1 \le p < \infty$, 对可测函数f定义*

$$\|f\|_p := (\int |f|^p d\mu)^{\frac{1}{p}},$$

称为f的p-范数. L^p空间定义为

$$L^p := \{f : \|f\|_p < \infty\}.$$

简言之, L^p就是由所有p方可积函数组成的空间.

在L^p中, 我们把两个几乎处处相等的函数认为是相等的. 因此, 严格地说, L^p的元素并不是可测函数, 而是由几乎处处相等的函数组成的等价类. 当然, 在实际运算时, 我们总是在每个这样的等价类中选取一个代表也就是一个p次可积的函数来进行.

为研究L^p的性质, 我们需要一点凸函数方面的知识.

定义10.1.2. 设I为$\mathbb{R}$中的区间, ϕ为I上的连续函数. 若$\forall x_1, x_2 \in I$, $\alpha_1, \alpha_2 > 0$, $\alpha_1 + \alpha_2 = 1$, 有

$$\phi(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 \phi(x_1) + \alpha_2(x_2), \tag{10.1.18}$$

或者等价地, 对$\forall n$, $\forall x_i \in I$, $i = 1, 2, \cdots, n$, $\alpha_i > 0$, $i = 1, 2, \cdots, n$, $\sum_{i=1}^n \alpha_i = 1$, 有

$$\phi(\sum_{i=1}^n \alpha_i x_i) \leq \sum_{i=1}^n \alpha_i \phi(x_i)), \tag{10.1.19}$$

则称φ为凸函数.

我们注意到, 10.1.18与10.1.19的等价性也即10.1.18蕴涵10.1.19可由归纳法导出.

下面给出一个凸函数的充分条件.

定理10.1.3. 记号同上. 设ϕ处处可导且ϕ'单调上升, 则ϕ为凸函数.

证明: 不妨设$x_1 < x_2$. 取数a, b使

$$\psi(x) := \phi(x) + ax + b, \quad x \in [x_1, x_2].$$

满足$\psi(x_1) = \psi(x_2) = 0$, 则10.1.18等价于在$[x_1, x_2]$上

$$\psi \leq 0. \tag{10.1.20}$$

设否, 则由于ψ连续, 故有一极大点$x_0 \in (x_1, x_2)$使$\psi(x_0) > 0$而$\psi'(x_0) = 0$. 由于$\psi' = \phi' + a$, 故ψ'单调上升, 因而在$[x_0, x_2]$上$\psi' \geq 0$, 这样就势必有$\psi(x_2) \geq \psi(x_0) > 0$, 矛盾. 完毕.

由此我们立即得到一个简单的判别准则.

命题10.1.4. 记号同上. 若ϕ处处二次可微且ϕ''处处非负, 则ϕ为凸函数.

这样, 我们就知道例如$\exp(t)$, t^p $(p \geq 1)$, $-\log(t)$都是凸函数.

我们需要的是如下的

引理10.1.5. 设$t_1, t_2 \geq 0$, $\alpha, \beta > 0$且$\alpha + \beta = 1$, $p \geq 1$. 则

$$t_1^\alpha t_2^\beta \leq \alpha t_1 + \beta t_2; \tag{10.1.21}$$

$$(\alpha t_1 + \beta t_2)^p \leq \alpha t_1 + \beta t_2. \tag{10.1.22}$$

证明: 10.1.21、10.1.22分别是由于$-\log(t)$、t^p是凸函数之故. 完毕.

借助于此引理, 我们可以证明两个重要的不等式.

定理10.1.6. *i) (Hölder不等式)设*$1 < p, q < \infty$*满足*$p^{-1} + q^{-1} = 1$, $f \in L^p$, $g \in L^q$, *则*$fg \in L^1$*且*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q;$$

ii) (Minkowvski不等式)设$f, g \in L^p$, *则*$f + g \in L^p$*且*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

证明: 如果f与g中有一个恒为0, 则两个不等式都是显然的, 因此我们假定他们都不为0.

由10.1.21有

$$|\frac{f(x)}{\|f\|_p}\frac{g(x)}{\|g\|_q}| \leq \frac{1}{p}|f(x)|^p\|f\|_p^p + \frac{1}{q}|f(x)|^q\|f\|_q^q,$$

两边取积分即得

$$\|f\|_p^{-1}\|g\|_q^{-1}\|fg\|_1 \leq \frac{1}{p} + \frac{1}{q} = 1,$$

i)得证.

由10.1.21有

$$\begin{aligned} |f(x) + g(x)|^p &= |\frac{\|f\|_p}{\|f\|_p + \|g\|_p}(f(x)\|f\|_p^{-1}) + \frac{\|g\|_p}{\|f\|_p + \|g\|_p}(g(x)\|g\|_p^{-1})|^p(\|f\|_p + \|g\|_p)^p \\ &\leq (\|f\|_p + \|g\|_p)^p(\frac{\|f\|_p}{\|f\|_p + \|g\|_p}(f(x)\|f\|_p^{-1})^p + \frac{\|g\|_p}{\|f\|_p + \|g\|_p}(g(x)\|g\|_p^{-1})^p), \end{aligned}$$

两边取积分即得

$$\|f + g\|_p^p \leq (\|f\|_p + \|g\|_p)^p,$$

再两边开p次方便得ii). 完毕.

我们给出几个简单的应用.

若$(X, \mathscr{F}, \mu)$是有限测度空间, 则对$p_1 \geq p_2 \geq 1$, 由Hölder不等式有:

$$\begin{aligned} &\int |f|^{p_2} d\mu = \int |f|^{p_2} 1 d\mu = \int |f|^{p_1\frac{p_2}{p_1}} 1 d\mu \\ \leq\ &(\int |f|^{p_1} d\mu)^{\frac{p_2}{p_1}} (\int 1 d\mu)^{\frac{p_1 - p_2}{p_1}} = (\mu(X))^{\frac{p_1 - p_2}{p_1}} (\int |f|^{p_1} d\mu)^{\frac{p_2}{p_1}}. \end{aligned}$$

因此我们得到

命题10.1.7. *若*$(X, \mathscr{F}, \mu)$*是有限测度空间*, $p_1 \geq p_2 \geq 1$, *则*$L^{p_1} \subset L^{p_2}$. *特别地, 当*$\mu(X) = 1$*时*, $p \mapsto \|f\|_p$*是单调上升的.*

注意在$\mu(X)=\infty$时此命题不成立, 例如对于$\mathbb{R}$上的lebesgue测度, $\frac{1}{|x|+1}\in L^2$但$\notin L^1$. 不过, 不管是否$\mu(X)=\infty$我们都有下面的结果:

定理10.1.8. *若$p_1>p_2\geq 1$, $f\in L^{p_1}\cap L^{p_2}$, 则$\forall p\in(p_2,p_1)$, $f\in L^p$.*

证明: 我们有

$$\int |f|^p=\int |f|^{p_1\frac{p-p_2}{p_1-p_2}}|f|^{p_2\frac{p_1-p}{p_1-p_2}}d\mu\leq (\int |f|^{p_1})^{\frac{p-p_2}{p_1-p_2}}(\int |f|^{p_2})^{\frac{p_1-p}{p_1-p_2}}.$$

完毕.

上面证明中出现的这个不等式即

$$\|f\|_p\leq \|f\|_{p_1}^{\frac{p_1(p-p_2)}{p(p_1-p_2)}}\|f\|_{p_2}^{\frac{p_2(p_1-p)}{p(p_1-p_2)}}$$

称为L^p间的插值不等式.

上面的Minkowski不等式可以有以下更一般的积分形式:

定理10.1.9. *设$(X,\mathscr{F},\mu)$与$(Y,\mathscr{G},\nu)$都是σ有限测度空间, $f(x,y)\in\mathscr{F}\times\mathscr{G}$. 则$\forall p\geq 1$,*

$$(\int_X|\int_Y f(x,y)\nu(dy)|^p\mu(dx))^{1/p}\leq \int_Y(\int_X |f(x,y)|^p\mu(dx))^{1/p}\nu(dy). \quad (10.1.23)$$

证明: 我们第一可假定f非负; 其次, 设$Y_n\uparrow Y$使$\nu(Y_n)<\infty$, 若10.1.23在以Y_n取代Y时成立的话, 则在两端取极限, 左边用单调收敛定理, 知10.1.23成立, 因此我们第二假定$\nu(Y)<\infty$. 最后, 若10.1.23对非负有界的可测函数成立的话, 由于对一般非负的f可找一列几乎处处收敛于它的非负有界的f_n, 则由单调收敛定理知对f也成立, 因此第三我们假定f是有界的.

对于非负有界可测函数f, 存在简单函数列f_n单调上升一致收敛到f. 而对简单函数, 10.1.23就是已知的Minkowski不等式, 所以有

$$\begin{aligned}(\int_X(\int_Y f_n(x,y)\nu(dy))^p\mu(dx))^{1/p} &\leq \int_Y(\int_X (f_n(x,y))^p\mu(dx))^{1/p}\nu(dy)\\ &\leq \int_Y(\int_X (f(x,y))^p\mu(dx))^{1/p}\nu(dy),\end{aligned}$$

取$n\to\infty$并用控制收敛定理便得到10.1.23. 完毕.

本定理可用一句话来记忆: 积分的范数小于等于范数的积分. 特别, 取$Y=\{1,2,\cdots\}$, $\mathscr{G}$为所有子集, ν为记数测度, 则得到:

推论10.1.10.

$$\int_X|\sum_{n=1}^{\infty}f_n(x)|^p\mu(dx)\leq (\sum_{n=1}^{\infty}(\int_X|f_n(x)|^p\mu(dx))^{1/p})^p$$

顺便说一下, 设μ是概率测度即$\mu(X)=1$, 在Hölder不等式中取$g=1$便得

$$(\int |f|d\mu)^p \le \int |f|^p d\mu.$$

这一不等式其实对一般的凸函数也成立, 即我们有:

定理10.1.11. *(Jensen不等式) 设$(X,\mathscr{F},\mu)$是概率测度, ϕ是凸函数, f可积. 则*

$$\phi(\int f d\mu) \le \int \phi(f) d\mu.$$

证明:

我们利用凸函数的以下性质:

$$\phi(t)-\phi(t_0) \ge \phi'_+(t_0)(t-t_0),$$

这里ϕ'_+表示右导数. 所以

$$\phi(f(x))-\phi(\int f d\mu) \ge \phi'_+(\int f d\mu)(f(x)-\int f d\mu),$$

两边取积分便有

$$\int \phi(f)d\mu - \phi(\int f d\mu) \ge 0.$$

完毕.

推论10.1.12. *设f非负可积, $E[f]=1$. 则*

$$E[fg] \le E[f\log f] + \log(E[e^g]).$$

证明(stroock: An introduction to the analysis of paths on a Riemannian manifold, p.162) . 不妨设f一致正且f,g均有界. 令

$$dQ := f dP.$$

则由Jensen不等式

$$\log(E^P[e^\psi]) = \log(E^Q[f^{-1}e^\psi]) \le -E^Q[\log f] + E^Q[\psi].$$

完毕.

我们知道,若f,g均二次可积,那么由Schwarz不等式, fg可积; Hölder不等式意味着可以将其中一个的可积性提高,另一个降低,这样两个相互补偿仍能得到fg的可积性; 上面的推论则告诉我们,这种补偿的范围还可以继续扩大: 如果一个指数可积,则另一个只要比可积稍微强一点, 即$f\log f$可积就行了.

此外,若没有$E[f]=1$的限制,则上述公式变为

$$E[fg] \le E[f]E[f\log f] + E[f]\log(E[e^g]).$$

我们再回到L^p空间. 我们要证明:

定理10.1.13. $\forall p\in[1,\infty)$, $(L^p,\|\cdot\|_p)$为*Banach*空间.

证明: 由Minkowski不等式, L^p为线性空间且$\|\cdot\|_p$为范数. 下面证明L^p是完备的.

设$f_n\in L^p$, $\lim_{m,n\to\infty}\|f_n-f_m\|_p=0$. 首先, 存在$n_1$使$m,n\geq n_1$时,

$$\|f_n-f_m\|_p<2^{-1};$$

接着, 存在$n_2>n_1$使$n,m\geq 2^{-2}$时

$$\|f_n-f_m\|_p<2^{-2};$$

一般地, n_k选好后, 可选$n_{k+1}>n_k$使$m,n>n_{k+1}$时

$$\|f_n-f_m\|_p<2^{-(k+1)}.$$

这样就得到一列$n_1,n_2,\cdots$. 显然有

$$\|f_{n_{k+1}}-f_{n_k}\|_p\leq 2^{-k}.$$

于是由推论10.1.10有

$$\int(|f_{n_1}|+\sum_{k=1}^{\infty}|f_{n_{k+1}}-f_{n_k}|)^p d\mu\leq(\|f_{n_1}\|_p+\sum_{k=1}^{\infty}\|f_{n_{k+1}}-f_{n_k}\|_p)^p<\infty.$$

因此

$$|f_{n_1}|+\sum_{k=1}^{\infty}|f_{n_{k+1}}-f_{n_k}|<\infty \text{ a.e.},$$

于是级数

$$f_{n_1}+\sum_{k=1}^{\infty}(f_{n_{k+1}}-f_{n_k})$$

几乎处处收敛. 记其极限函数为f, 则$\lim_{k\to\infty}f_{n_k}=f$ a.e. 且

$$\int|f|^p d\mu\leq\int(|f_{n_1}|+\sum_{k=1}^{\infty}|f_{n_{k+1}}-f_{n_k}|)^p d\mu<\infty,$$

因此$f\in L^p$. 再由Fatou引理有

$$\lim_{n\to\infty}\int|f-f_n|^p d\mu=\lim_{n\to\infty}\int\lim_{k\to infty}|f_{n_k}-f_n|^p d\mu\leq\lim_{n,n_k\to\infty}\int|f_{n_k}-f_n|^p d\mu=0,$$

即$\lim_{n\to\infty}\|f-f_n\|_p=0$. 完毕.

§10.2 L^∞空间

$p=\infty$时不可以通过积分定义范数. 此时我们定义

$$\|f\|_\infty=\inf_{\mu(E)=0}\sup_{x\in E^c}|f(x)|.$$

例如, 设$([0,1],\mathscr{B}([0,1]),dx)$上的函数$f$定义为

$$f(x)=1,\ 若x为有理数; 0,\ 若x为无理数,$$

则

$$\|f\|_\infty=0.$$

又若f定义为

$$f(x)=(1+x)^{-1}\ 若x>0; f(2)=2,$$

则$\|f\|_\infty=1$. 若$f(x)=x^{-1}$, 则$\|f\|_\infty=\infty$.

再定义

$$L^\infty:=\{f:\|f\|_\infty<\infty\}.$$

我们有

定理10.2.1. $\|\cdots\|_\infty$为L^∞上的范数. L^∞在此范数下为$Banach$空间.

证明: 先证明第一个结论. 首先我们注意到在$\|\cdots\|_\infty$的定义中, 下确界是可以达到的: 设

$$c=\inf_{\mu(E)=0}\sup_{x\in E^c}|f(x)|,$$

则$\forall n$, 存在E_n, $\mu(E_n)=0$, 使

$$c>\sup_{x\in E_n^c}|f(x)|-\frac{1}{n}.$$

于是, 令$E=\cup_n E_n$, 则$\mu(E)=0$且

$$c\geq\sup_{x\in E^c}|f(x)|,$$

故

$$c=\sup_{x\in E^c}|f(x)|.$$

这样, 设$\|f\|_\infty=\sup_{x\in E^c}|f(x)|$, $\|g\|_\infty=\sup_{x\in E^c}|g(x)|$, $\mu(E)=\mu(F)=0$, 则

$$\sup_{x\in(E\cup F)^c}\{|f(x)+g(x)|\}\leq\|f\|_\infty+\|g\|_\infty,$$

因而

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty.$$

现在证明L^∞的完备性. 设$f_n \in L^\infty$, $\lim_{m,n\to\infty}\|f_n - f_m\|_\infty = 0$. 再设

$$\|f_n\|_\infty = \sup_{x\in E_n^c} |f_n(x)|,$$

$$\|f_n - f_m\|_\infty = \sup_{x\in E_{m,n}^c} |f_m(x) - f_n(x)|,$$

其中$E_n, E_{m,n}$均是零测集. 令$E = (\cup_n E_n)\cup(\cup_{m,n}E_{m,n})$, 则$\mu(E)=0$, 因而

$$\|f_n\|_\infty = \sup_{x\in E^c} |f_n(x)|,$$

$$\|f_n - f_m\|_\infty = \sup_{x\in E^c} |f_m(x) - f_n(x)|.$$

定义

$$f(x) = \lim_{n\to\infty} f_n(x) \text{ 若} x\in E^c; 0, \text{ 若} x \in E,$$

则$f\in L^\infty$且$\|f - f_n\| \to \infty$. 完毕.

若$f, g\in L^2$, 定义

$$(f,g) := \int fgd\mu, \tag{10.2.24}$$

容易验证$(\cdot,\cdot)$是L^2中的内积, 且相应的范数就是$\|\cdot\|$. 因此我们有

定理10.2.2. *带有内积10.2.24的L^2是Hilbert空间.*

最后我们证明:

命题10.2.3. *若$(X,\mathscr{F},\mu)$是概率空间, 则$\forall f\in L^\infty$, 有*

$$\|f\|_p \uparrow \|f\|_\infty.$$

证明: $p\mapsto \|f\|_p$单调上升是已证的, $\|f\|_p \le \|f\|_\infty$是显然的. 又$\forall\varepsilon>0$, 有

$$\mu(|f(x)| \ge \|f\|_\infty - \varepsilon) > 0,$$

故

$$\|f\|_p \ge (\|f\|_\infty - \varepsilon)(\mu(|f(x)| \ge \|f\|_\infty - \varepsilon)^{1/p}.$$

所以

$$\lim_{p\to\infty}\|f\|_p \ge \|f\|_\infty - \varepsilon.$$

由ε的任意性即得结论. 完毕.

11 赋号测度

§11.1 定义及基本性质

我们在构造$\mathbb{R}$上的Lebesgue-Stieltjes测度的时候,出发点是单调上升的函数(见例3.7.2). 然而由于一般的有限变差函数均可表示为两个增函数之差,故同样的推理对有限变差函数依然有效,只不过最后得出来的集函数不再是非负的. 这样一种集函数我们称之为赋号测度. 即有

定义11.1.1. *设$(X, \mathscr{F})$为可测空间,则$\mathscr{F}$上的可加集函数称为赋号测度.*

注11.1.2. *由于运算$+\infty - (+\infty)$没有意义,故在$\pm\infty$中,赋号测度只能取到一个,否则将会出现无意义的等式*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

如无特别申明, 我们一般遵循下面的

$$\text{约定} \qquad \mu(E) < \infty,\ \forall A \in \mathscr{F}. \tag{11.1.25}$$

命题11.1.3. *1. 设$E, F \in \mathscr{F}$, $F \subset E$. 若$\mu(E) > -\infty$,则$\mu(F) > -\infty$.*

2. 设$E_n, E \in \mathscr{F}$, $E_n \uparrow E$, 则

$$\lim_{n\to\infty} \mu(E_n) = \mu(E)$$

3. 设$E_n, E \in \mathscr{F}$, $E_n \downarrow E$, 且$|\mu(E_1)| < \infty$, 则

$$\lim_{n\to\infty} \mu(E_n) = \mu(E)$$

证明:

1.由

$$\mu(E) = \mu(F) + \mu(E \setminus F)$$

直接得到.

2与3的证明类似于测度的情形.

例11.1.4. *在一般的σ有限测度空间$(X, \mathscr{F}, \mu)$上, 设$f \in L^1(\mu)$. 令*

$$\nu(A) := \int_A f d\mu, \quad \forall A \in \mathscr{F},$$

则ν是赋号测度.

§11.2 Jordan-Hahn分解

在$\mathbb{R}$上, 利用有限变差函数可表为两个增函数之差, 容易证明由有限变差函数产生的赋号测度能表示为两个测度之差. 这一结果对一般的赋号测度也成立. 证明这个结果的基本想法是将该测度的"正集"与"负集"分开. 所谓"正集"、"负集"是指:

定义11.2.1. *设$E \in \mathscr{F}$使得对任意$F \in \mathscr{F}$, 若*

$$F \subset E \Longrightarrow \mu(F) \geq 0;$$

则E称为μ的正集.

引理11.2.2. *设$E \in \mathscr{F}$且$\mu(E) > 0$, 则存在正集$F \subset E$且$\mu(F) \geq \mu(E)$.*

证明: 第一步: 若存在$G \subset E$使得$\mu(G) < -1$, 则记$A_{11} = G$. 否则第一步终止.
若存在$G \subset E \setminus A_{11}$使得$\mu(G) < -1$, 则记$A_{12} = G$. 否则第一步终止.
继续下去, 这第一步一定会在进行有限步(设为n_1步)后终止. 记

$$A_1 := \sum_{k=1}^{n_1} A_{1k}.$$

第二步: 以$-\frac{1}{2}$代替上面的-1做同样的事情,得到A_2;
一般地, 以$-\frac{1}{n}$代替上面的-1做同样的事情,得到A_n.
令

$$F := E - \sum_{n=1}^{\infty} A_n.$$

则显然F满足要求.

定理11.2.3. *设$(X, \mathscr{F}, \mu)$为赋号测度空间,则存在$E_+ \in \mathscr{F}$, E为μ的正集且$\forall F \in \mathscr{F}$, $F \subset E^c$, 有$\mu(F) \leq 0$. E_+在下述意义下唯一: 若E_1, E_2均满足上述要求, 则*

$$\mu(E_1 \triangle E_2) = 0.$$

该E_+称为μ的正部, $E_- := E_+^c$称为负部.

证明: 唯一性显然,只需证存在性.
若$\forall F \in \mathscr{F}$, $\mu(F) \leq 0$, 则取$E_+ = \emptyset$.
下面设$\sup_{F \in \mathscr{F}} \mu(F) = a > 0$. 取$A_n \in \mathscr{F}$,使

$$\mu(A_n) \uparrow a.$$

由上一引理, 可设A_n均为正集. 令

$$E := \cup_{n=1}^{\infty} A_n.$$

则由于A_n都是正集, 故$\mu(E) = a$. 于是$a < \infty$.

该集合满足要求. 否则或者存在$E' \subset E_+$, $E' \in \mathscr{F}$使得$\mu(E') < 0$; 或者存在$E' \subset E_+^c$, $E' \in \mathscr{F}$使得$\mu(E') > 0$. 在前者,有

$$\mu(E_+ \setminus E') = \mu(E) - \mu(E') > \mu(E);$$

在后者,有

$$\mu(E_+ + E') = \mu(E) + \mu(E') > \mu(E).$$

两者均不可能.

Q.E.D.

继续上定理得记号, 令

$$\mu^+(A) := \mu(A \cap E), \quad \mu^-(A) := -\mu(A \cap E^c)$$

则得到μ的所谓的Jordan分解

$$\mu = \mu^+ - \mu^-,$$

μ^+, μ^-分别称为μ的上、下变差, $|\mu| := \mu^+ + \mu^-$称为μ的全变差, 而

$$X = E + E^c$$

则称为μ的Hahn分解. 若$f \in L^1(|\mu|)$, 则定义

$$\int_X f d\mu = \int_X f d\mu^+ - \int_X f d\mu^-.$$

§11.3 Radon-Nikodym定理

设$(X, \mathscr{F}, \mu)$为σ有限测度空间, ν是$\mathscr{F}$上的赋号测度且$|\nu|\sigma$有限. 若$\forall E \in \mathscr{F}$, 有

$$\mu(E) = 0 \Longrightarrow \nu(E) = 0,$$

则称ν关于μ绝对连续,记为$\nu \ll \mu$.

这一条件等价于:

$$\mu(E) = 0 \Longrightarrow |\nu|(E) = 0.$$

事实上, 若前者成立, 设$X = E_+ + E_-$为ν的Hahn分解, 则由$\mu(E) = \mu(E \cap E_+) + \mu(E \cap E_-)$ 知

$$\mu(E \cap E_+) = \mu(E \cap E_-) = 0.$$

因而

$$\nu(E \cap E_+) = \nu(E \cap E_-) = 0,$$

即

$$\nu^+(E) = \nu^-(E) = 0,$$

从而

$$|\nu|(E) = 0.$$

由积分的性质,例11.1.4中的集函数ν是关于μ绝对连续的赋号测度. 我们下面将证明, 本质上这种形式的赋号测度已囊括所有的关于μ绝对连续的赋号测度, 这就是

定理11.3.1. *(Radon-Nikodym定理)设$(X, \mathscr{F}, \mu)$为σ有限测度空间, ν为$\mathscr{F}$上的σ有限赋号测度. 若$\nu << \mu$,则存在一关于$\mu\sigma$可积的函数f,使得*

$$\nu(A) = \int_A f d\mu, \quad \forall A \in \mathscr{F}.$$

若ν有限,则$f \in L^1(d\mu)$.

该函数f在几乎处处的意义下是唯一的, 记为$f = \frac{d\nu}{d\mu}$.

唯一性显然. 为证明存在性我们需要下面的预备结果.

引理11.3.2. *条件同上且设$\nu \neq 0$. 则存在$n \in \mathbb{N}^+$及$B \in \mathscr{F}$使得$\mu(B) > 0$ 且B是$\nu - \frac{1}{n}\mu$的正集.*

证明:设$X = A_n^+ + A_n^-$是$\nu - \frac{1}{n}\mu$的Hahn分解. 令

$$A_0^+ = \cup_{n=1}^{\infty} A_n^+, \quad A_0^- = \cap_{n=1}^{\infty} A_n^1.$$

则$A_0^+ = X - A_0^-$且

$$\nu(A_0^-) \leq \frac{1}{n}\mu(A_0^-), \quad \forall n.$$

于是$\nu(A_0^-) = 0$. 因而$\nu(A_0^+) > 0$. 故存在n使得$\nu(A_n^+) > 0$. 再由绝对连续性, $\mu(A_n^+) > 0$. 因此该n和$B = A_n^+$满足要求.

存在性的证明: 由Jordan分解,$\nu = \nu_+ - \nu_-$,因而只需要对ν_+, ν_-分别证明.这样我们将假定ν是非负的; 再由于我们可以先在μ与ν均有限的区域上考虑,然后把所得的函数拼起来,故又可以假设他们均是有限测度. 令

$$\mathscr{L} := \{f : \ f \geq 0, f \in L^1(\mu), \int_A f du \leq \nu(A), \ \forall A \in \mathscr{F}\}.$$

则$\mathscr{L}$非空(它包含恒为零的函数). 此外,设$f, g \in \mathscr{L}$, 令$h := f \vee g$, 则$\forall A \in \mathscr{F}$,

$$\begin{aligned}\int_A h d\mu &= \int_{A\cap\{f<g\}} h d\mu + \int_{A\cap\{f<g\}} h d\mu \\ &= \int_{A\cap\{f<g\}} g d\mu + \int_{A\cap\{f<g\}} f d\mu \\ &\leq \nu(A \cap \{f < g\}) + \nu(A \cap \{f < g\}) = \nu(A).\end{aligned}$$

故$h \in \mathscr{L}$.

令

$$a := \sup\{\int_X f d\mu : f \in \mathscr{L}\}.$$

取$f_n \in \mathscr{L}$使得

$$\lim_{n\to\infty} \int f_n d\mu = a.$$

必要时以$f_1 \vee f_2 \vee \cdots f_n$代替$f_n$, 可假定$\{f_n\}$单调上升. 令

$$f(x) := \lim_{n\to\infty} f_n(x).$$

则由单调收敛定理知

$$\int_X f d\mu = \lim_{n\to\infty} f_n d\mu = a.$$

往证

$$\nu(E) = \int_E f d\mu, \quad \forall E \in \mathscr{F}.$$

令

$$\lambda(E) := \nu(E) - \int_E f d\mu.$$

则λ为有限测度且$\lambda \ll \mu$. 若$\lambda \neq 0$, 则根据引理, 就存在$\varepsilon > 0$及$B \in \mathscr{F}$使得$\mu(B) > 0$且对任意$E \in \mathscr{F}$有

$$\varepsilon\mu(E \cap B) \leq \lambda(E \cap B).$$

于是,令

$$h = f + \varepsilon 1_B.$$

则对任意$E \in \mathscr{F}$

$$\begin{aligned}\int_E h d\mu &= \int_E f d\mu + \varepsilon\mu(E \cap B) \\ &\leq \int_{E\setminus B} f d\mu + \nu(E \cap B) \\ &\leq \nu(E).\end{aligned}$$

因此$h \in \mathscr{L}$. 但

$$\int_X h d\mu = \int_X f d\mu + \varepsilon\mu(B) > a,$$

这与a的定义矛盾. 这矛盾说明$\lambda = 0$. 所以

$$\nu(E) = \int_E f d\mu, \ \ \forall E \in \mathscr{F}.$$

. 证毕.

我们在这里使用了记号$f = \frac{d\nu}{d\mu}$, 但这是一整体记号,在目前看来与任何形式的商都没有什么关系. 不过,至少在离散情形,它确实是一个商;而在相当多的一般情形,它是某类商的极限, 就像普通的微分一样. 但要讲清这点,就要用到鞅的知识, 超出了本课程的范围.

现在我们所能做的,是可以说明形式上我们可以将它当成商运算.

命题11.3.3. *设μ, ν分别是$(X, \mathscr{F})$上的σ有限的测度和赋号测度, $\nu \ll \mu$. 则对于任意非负可测函数f, 有*

$$\int f d\nu = \int f \frac{d\nu}{d\mu} d\mu.$$

证明: 函数形式的单调类定理.

推论11.3.4. *设μ, ν是$(X, \mathscr{F})$上的σ有限测度, λ是其上的σ有限赋号测度. $\lambda \ll \nu \ll \mu$. 则*

$$\frac{d\lambda}{d\mu} = \frac{d\lambda}{d\nu}\frac{d\nu}{d\mu}.$$

12 条件期望

我们回忆一下本科教材中数学期望的概念: 设ξ是概率空间$(X, \mathscr{F}, \mu)$上的随机变量, F是ξ的分布函数, 则ξ的数学期望定义为

$$E\xi = \int_{-\infty}^{\infty} x dF(x).$$

利用变量代换公式, 上式可写为

$$E\xi = \int \xi d\mu.$$

以后, 在概率空间上积分时, 我们就用E代替积分符号.

我们换一个观点看这个问题. 直观上说, 数学期望应该是最接近于原随机变量的常数. 这个直观感觉可以从数学上得到严格证明——至少对于平方可积的随机变量是如此. 实际上, 设$\xi \in L^2$, 令

$$f(c) = E[|\xi - c|^2].$$

则易知$c = E\xi$时$f(c)$达到最小值.

我们知道, 若以$\mathscr{F}_0$表示X上的平凡σ代数, 即

$$\mathscr{F}_0 := \{X, \emptyset\},$$

则当且仅当一个函数(几乎处处)为常数时, 它关于$\mathscr{F}_0$可测. 由于$L^2(X, \mathscr{F}_0, \mu)$是$L^2(X, \mathscr{F}, \mu)$的闭子空间, 因此用泛函分析的语言来说, $L^2(X, \mathscr{F}, \mu)$中函数的数学期望就是它在$L^2(X, \mathscr{F}_0, \mu)$上的正交投影.

在刻画实际问题时, σ代数表示信息, 这样$\mathscr{F}$便表示全体有关的信息, 而平凡的$\mathscr{F}_0$表示我们不掌握任何相关的信息. 因此, 当我们要求期望时, 我们实际上是在没有任何(有利)条件下寻求ξ的最佳近似.

而当我们有一些信息－－我们用σ代数$\mathscr{G}$表示这些信息全体－－可用时, 我们理应获得更好的近似. 这一想法在数学上引导出了条件数学期望或简称条件期望的概念.

§12.1 L^2中函数的条件期望

设$\mathscr{G}$是$\mathscr{F}$的子σ代数. 简记

$$L^p(\mathscr{F}) := L^2(X, \mathscr{F}, \mu); \quad L^p(\mathscr{G}) := L^2(X, \mathscr{G}, \mu).$$

我们先证明一个关于一般的p的结论.

引理12.1.1. $L^p(\mathscr{G})$是$L^p(\mathscr{F})$的闭子空间.

证明: 显然$L^p(\mathscr{G}) \subset L^p(\mathscr{F})$. 设$f_n \in L^(\mathscr{G})$, $f \in L^p(\mathscr{F})$而$\lim_{n\to\infty} \|f_n - f\|_{L^p(\mathscr{F})} = 0$, 则

$$\lim_{m,n\to\infty} \|f_m - f_n\|_{L^p(\mathscr{G})} = \lim_{m,n\to\infty} \|f_m - f_n\|_{L^p(\mathscr{F})} = 0.$$

由于$L^p(\mathscr{G})$是完备的, 故存在$\tilde{f} \in \mathcal{L}^2(\mathscr{G})$使$\lim_{n\to\infty} \|f_n - \tilde{f}\|_{L^p(\mathscr{G})} = 0$. 显然$f = \tilde{f}$ a.e., 因此$f \in L^p(\mathscr{G})$. 完毕.

于是特别地, $L^2(\mathscr{G})$是$L^2(\mathscr{F})$的闭子空间. 将后者到前者的正交投影记为$P_{\mathscr{F},\mathscr{G}}$, 我们给出:

定义12.1.2. 设$f \in L^2(\mathscr{F})$, 则f关于$\mathscr{G}$的条件期望定义为

$$E(f|\mathscr{G}) := P_{\mathscr{F},\mathscr{G}} f.$$

我们有

定理12.1.3. 设$f \in L^2(\mathscr{F})$, $g \in L^2(\mathscr{G})$, 则下列条件等价:

1) $g = E[f|\mathscr{G}]$;

2) $\forall h \in L^2(\mathscr{G})$, $\int fh\ d\mu = \int gh\ d\mu$;

3) $\forall h \in L^\infty(\mathscr{G})$, $\int fh\ d\mu = \int gh\ d\mu$;

4) $\forall E \in \mathscr{G}$, $\int_E f\ d\mu = \int_E g\ d\mu$.

证明: 首先, 由正交投影的性质, $1^o \Longleftrightarrow 2^o$.

又显然$2^o \Longrightarrow 3^o \Longrightarrow 4^o$, 因此只要证明$4^o \Longrightarrow 2^o$.

设4^o成立, 则由积分的线性性知对任意简单函数$h = \sum_{i=1}^n a_i 1_{E_i}$, $E_i \in \mathscr{G}$, 有

$$\int fh\ d\mu = \int gh\ d\mu.$$

而对于一般的$h \in L^2(\mathscr{G})$, 均可以找到这样的简单函数列h_n使$h_n \to h$ a.e.且$|h_n| \leq |h|$, 因此由控制收敛定理知上式对任意$h \in L^2(\mathscr{G})$成立. 完毕.

条件期望有如下的基本性质:

命题12.1.4. 1^0, $E[af + bg|\mathscr{G}] = aE[f|\mathscr{G}] + bE[g|\mathscr{G}]$.

2^0 设$\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{F}$, 则$E[E[f|\mathscr{G}_2]|\mathscr{G}_1] = E[f|\mathscr{G}_1]$. 特别地, $\forall \mathscr{G} \subset \mathscr{F}$, $E[E[f|\mathscr{G}]] = E[f]$.

3^0 $\forall h \in L^\infty(\mathscr{G})$, $E[fh|\mathscr{G}] = hE[f|\mathscr{G}]$.

4^0 $\forall f \in L^2(\mathscr{F})$, $f \geq 0$, $E[f|\mathscr{G}] \geq 0$.

5^0 $f \geq g \Longrightarrow E[f|\mathscr{G}] \geq E[g|\mathscr{G}]$

6^0 $|E[f|\mathscr{G}]| \leq E[|f||\mathscr{G}]$.

证明:

1^0 由正交投影的线性性.

2^0 $\mathscr{G}_1 \subset \mathscr{G}_2$时, $L^2(\mathscr{G}_1) \subset L^2(\mathscr{G}_2)$, 因此$P_{G_1} = P_{G_1} P_{G_2}$.

3^0 $\forall E \in \mathscr{G}$,

$$\int_E hE[f|\mathscr{G}]d\mu = \int 1_E hE[f|G]d\mu = \int E[f1_E h|G]d\mu = \int f1_E h d\mu = \int_E fhd\mu.$$

4^0 $\forall E \in \mathscr{G}$, $\int E[f|\mathscr{G}]d\mu = \int fd\mu \geq 0$, 所以$E[f|\mathscr{G}] \geq 0$.

5^0 由4^0; 6^0由5^0. 完毕.

§12.2 L^1中函数的条件期望

由上面的性质6^0, 可将$E[f|\mathscr{G}]$的定义域从$L^2(\mathscr{F})$扩充到$L^1(\mathscr{F})$. 具体地说, 我们有

定理12.2.1. $f \mapsto E[f|\mathscr{G}]$可唯一扩充为$L^1(\mathscr{F})$到$L^1(\mathscr{G})$的连续线性映射.

证明: 设$f\in L^1(\mathscr{F})$, 令$f_n:=(f\vee(-n))\wedge n$, 则$f_n\in L^2(\mathscr{F})$. 于是

$$|E[f_n-f_m|\mathscr{G}]|\le E[|f_n-f_m||\mathscr{G}],$$

所以

$$E|E[f_n|\mathscr{G}]-E[f_m|\mathscr{G}]|\le E|f_n-f_m|\le(E|f_n-f_m|^2)^{1/2}\ \text{Hölder不等式}.$$

但由控制收敛定理

$$\lim_{n\to\infty}\|f_n-f\|_2=0,$$

因而

$$\lim_{n\to\infty}\|E[f_n|\mathscr{G}]-E[f_m|\mathscr{G}]\|_1=0,$$

于是由$L^1(\mathscr{G})$的完备性, 知$E[f_n|\mathscr{G}]$在L^1中收敛. 我们于是定义

$$E[f|\mathscr{G}]:=\lim_{n\to\infty}E[f_n|\mathscr{G}].$$

自然, 这样定义的$E[f|\mathscr{G}]\in L^1(\mathscr{G})$. 下面证明连续性. 我们有

$$E|E[f|\mathscr{G}]-E[g|\mathscr{G}]|=E\lim_{n\to\infty}|E[f_n|\mathscr{G}]-E[g_n|\mathscr{G}]|\le\lim_{n\to\infty}E|f_n-g_n|\le\lim_{n\to\infty}E|f_n-g_n|=\|f-g\|_1,$$

因此$f\mapsto E[f|\mathscr{G}]$是$L^1(\mathscr{F})$到$L^1(\mathscr{G})$的连续映射.

再证明该映射是线性的. 设$f,g\in L^1(\mathscr{F})$, $\alpha,\beta\in\mathbb{R}$, 则可取$f_n,g_n\in L^2(\mathscr{F})$, $\lim_{n\to\infty}\|f-f_n\|_1=0$, $\lim_{n\to\infty}\|g-g_n\|_1=0$, 于是又有$\lim_{n\to\infty}\|\alpha f+\beta g-(\alpha f_n+\beta g_n)\|_1=0$. 故有

$$E[\alpha f+\beta g|\mathscr{G}]=\lim_{n\to\infty}E[\alpha f_n+\beta g_n|\mathscr{G}]=\lim_{n\to\infty}(\alpha E[f_n|\mathscr{G}]+\beta E[g_n|\mathscr{G}])=\alpha E[f|\mathscr{G}]+\beta E[g|\mathscr{G}]$$

我们有类似于定理12.1.3的

定理12.2.2. *设$f\in L^1(\mathscr{F})$, 则下列条件等价:*

1) $g=E[f|\mathscr{G}]$.

2) $\int_A gd\mu=\int_A fd\mu,\ \forall A\in\mathscr{G}$.

3) $\forall h\in L^\infty(\mathscr{G}),\ \int fhd\mu=\int ghd\mu$.

证明: 1)$\Longrightarrow$ 2):因为$E[f|\mathscr{G}]=\lim_n E[f_n|\mathscr{G}]\ (L^1)$, 所以

$$\int_A E[f|\mathscr{G}]d\mu=\lim_{n\to\infty}\int_A E[f_n|\mathscr{G}]d\mu=\lim_{n\to\infty}\int_A f_nd\mu=\int_A fd\mu.$$

2)$\Longrightarrow$ 3): 显然.

3)$\Longrightarrow$ 1): 由3), 有

$$\forall h\in L^\infty(\mathscr{G}),\ \int E[f|\mathscr{G}]hd\mu=\int ghd\mu,$$

所以$E[f|\mathscr{G}]=g$. 完毕.

通过L^2中的函数过渡, 同样可证明:

命题12.2.3. 1^0 设$\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{F}$, 则$E[E[f|\mathscr{G}_2]|\mathscr{G}_1] = E[f|\mathscr{G}_1]$. 特别地, $\forall \mathscr{G} \subset \mathscr{F}$, $E[E[f|\mathscr{G}]] = E[f]$.

2^0 $\forall h \in L^\infty(\mathscr{G})$, $E[fh|\mathscr{G}] = hE[f|\mathscr{G}]$.

3^0 $\forall f \in L^1(\mathscr{F})$, $f \geq 0$, $E[f|\mathscr{G}] \geq 0$.

5^0 $f \geq g \Longrightarrow E[f|\mathscr{G}] \geq E[g|\mathscr{G}]$

6^0 $|E[f|\mathscr{G}]| \leq E[|f||\mathscr{G}]$.

同期望一样, 条件期望也有Hölder、Minkowski和Jensen不等式. 其证明也是类似的, 读者可自己模仿一下.

命题12.2.4. *i)若φ是凸函数, 则*

$$\phi(E[f|\mathscr{G}]) \leq E[\phi(f)|\mathscr{G}].$$

特别, 若$f \in L^p(\mathscr{F})$, $1 \leq p \leq \infty$, 则

$$\|E[f|\mathscr{G}]\|_p \leq \|f\|_p.$$

ii)若$f \in L^p$, $g \in L^q$, $1 < p, q < \infty$, $p^{-1} + q^{-1} = 1$, 则

$$E[|fg||\mathscr{G}] \leq (E[|f|^p|\mathscr{G}])^{1/p}(E[|g|^q|\mathscr{G}])^{1/q}.$$

iii)若$f, g \in L^p$, 则

$$(E[|f+g|^p|\mathscr{G}])^{1/p} \leq (E[|f|^p|\mathscr{G}])^{1/p} + (E[|g|^p|\mathscr{G}])^{1/p}.$$

习题: 设$\mathscr{G}_1, \mathscr{G}_2$均是$\mathscr{F}$的子$\sigma$代数且无区别(即$\forall A_1 \in \mathscr{G}_1$ 存在$A_2 \in \mathscr{G}_2$, 使$\mu(A_1 \triangle A_2) = 0$), 则

$$E[f|\mathscr{G}_1] = E[f|\mathscr{G}_2].$$

§12.3 有限σ代数时条件期望的计算

虽然我们已给出了条件期望的定义, 也知道了它的一些性质, 但要实际算出一个条件期望来却远远不是一件简单的事情. 尽管如此, 我们还是希望能在最简单的情况下给出条件期望的算法.

最简单的情况自然是有限σ代数的情况, 即$\mathscr{G}$只包有有限个元素. 在本小节中我们恒做这个假定.

为搞清有限σ代数的结构, 我们引进原子的概念. 这是一个十分简单的概念, 远不如原子弹可怕.

定义12.3.1. *设$E \in \mathscr{G}$, $E \neq \emptyset$. 若在E的所有真子集中, 只有$\emptyset \in \mathscr{G}$, 则$E$称为$\mathscr{G}$的原子.*

我们有:

命题12.3.2. 设$\mathscr{G}$中每个非空集合均为它所包含的原子的并集.

证明: 设$A \in \mathscr{G}$, 首先证明A必包含一个原子. 若A本身是原子, 则定理已证; 若A不是原子, 存在A的真子集$A_1 \in \mathscr{G}$. 同样地, A_1要么是原子, 此时定理已证; 要么A_1有一个真子集$A_2 \in \mathscr{G}$; 把这一程序继续下去, 我们就可找到一列$A_1, A_2, \cdots$. 由于$\mathscr{G}$只有有限个元素, 而诸A_i全是$\mathscr{G}$的互不相同的元素, 故这一程序必定会在有限步——设是第r步——时终止. 这样, A_r就必须是原子.

现在, 设$B_1, \cdots, B_m$是含于A的所有原子, 令

$$B = \cup_{k=1}^m B_k.$$

则$B \subset A$. 若$A - B \neq \emptyset$, 则由前述道理, 在$A - B$中必然存在一个原子, 而这与$B_1, \cdots, B_m$是含于A的所有原子矛盾, 因此$A = B$. 完毕.

现在可以计算条件期望了. 设$\mathscr{G}$共有n个概率为正的原子, 他们是$A_1, \cdots, \mathcal{A}_n$. 我们有

定理12.3.3. 设$f \in L^1(\mathscr{F})$, 则

$$E[f|\mathscr{G}] = \sum_{i=1}^n \alpha_i 1_{A_i}, \text{ 其中} \alpha_i := \frac{1}{P(A_i)} E[f 1_{A_i}].$$

证明: 由于$A_i \in \mathscr{G}$, 所以$\sum_{i=1}^n \alpha_i 1_{A_i} \in \mathscr{G}$, 于是只要证明$\forall A \in \mathscr{G}$,

$$\int_A \sum_{i=1}^n \alpha_i 1_{A_i} dP = \int_A f dP.$$

由命题12.3.2, 又只需证明A是原子时上式成立. 但$A = A_i$时, 上式两边的值显然都是$\int_{A_i} f d\mu$. 完毕.

§12.4 给定随机变量时的条件期望

设f是随机变量, 以$\sigma(f)$记f产生的σ代数, 即:

$$\sigma(f) := \{f^{-1}(B), B \in \mathscr{B}(\mathbb{R})\}.$$

设$g \in L^1(\mathscr{F})$, 定义给定f时的条件期望为

$$E[g|f] = E[g|\sigma(f)].$$

由定理6.2.3, 存在可测映射φ使

$$E[g|f] = \varphi(f),$$

这启发我们定义“已知$f=r$时，g的条件期望”为

$$E[g|f=r]=\varphi(r).$$

我们于是有:

定理12.4.1. 设$g\in L^1(\mathscr{F})$, f的分布函数为F, 则

$$E[g]=\int_{-\infty}^{\infty}E[g|f=r]dF(r).$$

证明:

$$E[g]=E[E[g|f]]=E[\varphi(f)]=\int_{-\infty}^{\infty}\varphi(r)dF(r)=\int_{-\infty}^{\infty}E[g|f=r]dF(r).$$

完毕.

我们注意到, 若f只取有限个值, 即存在$r_i\in\mathbb{R}, A_i>0,\ i=1,\cdots,n$, 诸$A_i$不交, 使

$$f=\sum_{i=1}^{n}r_i 1_{A_i}$$

而

$$X=\cup_{i=1}^{n}A_i,\quad \mu(A_i)>0$$

则$\sigma(f)$是有限σ代数, 于是

$$E[g|f=r_i]=\frac{1}{\mu(A_i)}E[g1_{A_i}].$$

又dF只在诸r_i处有负荷$\mu(A_i)$, 所以

$$E[g]=\sum_{i=1}^{n}\frac{1}{\mu(A_i)}E[g1_{A_i}]\mu(A_i).$$

这样我们就回到了已经在初等概率论中熟悉了公式. 特别当$g=1_A, A\in\mathscr{F}$时, 我们回到了全概率公式.

为加深对这一公式的理解, 我们来看一个简单的例子. 设一个学校共有n个学生(用ω_i表示之), 分为m个系(自然$n>m$). 我们给这些系(数学、机械、$\cdots$,)编号为$1,2,\cdots,m$, 第i系的人数为n_i (于是$\sum_{i=1}^{m}n_i=n$). 设f定义为

$$f(\omega_i)=j,\ 若\omega_i\in 第j系.$$

设g为一与学生有关的数字(成绩、身高、$\cdots$), μ为概率:

$$\mu(\omega_i)=\frac{1}{n},$$

则

$$E[g|f=j]=\frac{1}{\mu(\text{第}j\text{系})}E[g1_{\text{第}j\text{系}}]=\frac{n}{n_j}\sum_{\omega_i\in\text{第}j\text{系}}g(\omega_i)\frac{1}{n}$$

$$=\sum_{\omega_i\in\text{第}j\text{系}}g(\omega_i)\frac{1}{n_j}=\text{ 第}j\text{系学生的平均水平}$$

$$\text{全体学生的平均水平}=E[g]=\sum_{j=1}^{m}\text{ 第}j\text{系学生的平均水平}\times\frac{n_j}{n}.$$

因此, 学校在计算全校学生的平均水平时, 并不需要傻算, 而只需将每系的平均水平再加权平均就行了.

§12.5 条件概率

设$A\in\mathscr{F}$, 给定$\mathscr{G}$时A发生的条件概率定义为

$$P(A|\mathscr{G}):=E[1_A|\mathscr{G}].$$

既然条件期望涉及到"期望"二字, 我们自然会想到它会不会是关于某种"条件概率"的积分呢? 也就是说, 是否存在一个二元函数$P(\omega,A),\omega\in\Omega,A\in\mathscr{F}$, 使得$\forall\omega$, $A\to P(\omega,A)$为概率测度, 而$\forall A\in\mathscr{F}$, $\omega\to P(\omega,A)\in\mathscr{G}$, 且对任意$f\in L^1(\mathscr{F})$有

$$E[f|\mathscr{G}](\omega)=\int f(\omega')P(\omega,d\omega')?$$

上述的$P(A|\mathscr{G})$自然是一个侯选对象. 但这里必须非常小心. 例如若$A,B\in\mathscr{F}$且不交, 虽然有

$$P(A\cup B|\mathscr{G})=P(A|\mathscr{G})+P(B|\mathscr{G})\text{ a.s.},$$

但其中的例外集一般是依赖于A、B的. 所以只有当$\mathscr{F}$的元素为可数个时, 我们才可以立即断言可找出公共的例外集. 由于一般说来$\mathscr{F}$ 很难保证只有可数个元素, 所以我们并不能轻易地下结论这样的$P(\omega,A)$存在. 实际上, 一般说来它是不存在的. 但我们有下面的定理:

定理12.5.1. *若Ω是Polish空间, $\mathscr{F}$是其Borelσ代数(由所有开集产生的σ代数), $\mathscr{G}$是$\mathscr{F}$的子σ代数. 则存在存在一个二元函数$P(\omega,A),\omega\in\Omega,A\in\mathscr{F}$, 使得$\forall\omega$, $A\to P(\omega,A)$为概率测度, 而$\forall A\in\mathscr{F}$, $\omega\to P(\omega,A)\in\mathscr{G}$, 且对任意$f\in L^1(\mathscr{F})$有*

$$E[f|\mathscr{G}](\omega)=\int f(\omega')P(\omega,d\omega')$$

$P(\cdot,\cdot)$称为关于$\mathscr{G}$的正则条件概率.

为了证明这个定理,也为了下一节讨论概率测度的弱收敛的需要,我们先研究一下Polish空间上概率测度的几个基本性质.

设S为Polish空间, $\mathscr{S}$是其上的Borel代数.

定义12.5.2. *设P是$(S,\mathscr{S})$上的概率测度,则P是正则的, 即对任意$A\in\mathscr{S}$,*

$$P(A) = \sup\{P(K),\ K\subset A \text{ 且}K\text{紧}\} \tag{12.5.26}$$
$$= \inf\{P(O),\ O\supset A \text{ 且}O\text{开}\} \tag{12.5.27}$$

证明: 先设$A=S$. 设$\{x_1,x_2,\cdots\}$为S的稠密子集, 令

$$A_{nk}:=\{x:\rho(x,x_k)<\frac{1}{n}\}.$$

则

$$\cup_k A_{nk}=S,\quad \forall n.$$

设$\varepsilon>0$. 对每一n, 选k_n使

$$P(\cup_{k=1}^{k_n}A_{nk})>1-\varepsilon 2^{-k}.$$

令

$$B:=\cap_n\cup_{k=1}^{k_n}A_{nk}.$$

则B是完全有界集,因而由S的完备性,$K:=\bar{B}$是紧集且

$$P(K^c)\le\sum_n \varepsilon 2^{-n}=\varepsilon.$$

现设A是闭集,则$K\cap A$是紧集且

$$P(A)-P(K\cap A)\le P(K^c)<\varepsilon.$$

而如果令

$$A_n:=\{x:\rho(x,A)<\frac{1}{n}\}.$$

则$A_n\downarrow A$, 所以$P(A_n)\downarrow P(A)$.

于是,令$\mathscr{E}$表示$\mathscr{S}$中满足(12.5.26)的集合全体，则$\mathscr{E}$包含了所有闭集. 容易验证$\mathscr{E}$为λ-类, 因此由$\lambda-\pi$定理, $\mathscr{E}=\mathscr{S}$.

注12.5.3. *这里完备性不能去掉, 见Billingsley.*

为什么要研究测度的正则性呢? 这与下面的考虑有关.

设$\mathscr{A}_0$是代数, $\mathscr{A}_1$是一紧集族, μ是定义在$\mathscr{A}_i$, $i=0,1$所产生的代数$\mathscr{A}$上的的非负有限可加的有限集函数. 由测度扩张定理, μ可扩张到$\sigma(\mathscr{A}_0)$上去的充要条件是它在$\mathscr{A}_0$上是可列可加的. 而这个可列可加性可以通过下面的结果得到.

引理12.5.4. 若对任意$A \in \mathscr{A}_0$,

$$\mu(A) \quad = \quad \sup\{\mu(K),\ K \subset A \text{ 且} K \in \mathscr{A}_1\}$$

则μ在$\mathscr{A}_0$上是可列可加的.

证明: 下面的证明的思想和构造Lebesgue-Stieltjes测度的思想是一样的,即用紧集的有限覆盖性质.

令

$$\mathscr{A}_2 := \{A : A^c \in \mathscr{A}_1\}.$$

则$\mathscr{A}_2$为一开集族且$\mathscr{A}_2 \subset \mathscr{A}$, 因而$\mu$在$\mathscr{A}_2$上有定义. 由于$\mathscr{A}_0$是代数, 所以

$$\mu(A) \quad = \quad \inf\{\mu(O),\ O \supset A \text{ 且} O \in \mathscr{A}_2\}, \quad \forall A \in \mathscr{A}_0.$$

设$A, A_n \in \mathscr{A}_0$且

$$A = \sum_{n=1}^{\infty} A_n.$$

则对任意N,

$$\sum_{n=1}^{N} A_n \subset A$$

故

$$\sum_{n=1}^{N} \mu(A_n) \leq \mu(A).$$

所以

$$\sum_{n=1}^{\infty} \mu(A_n) \leq \mu(A).$$

下面证明反向不等式. 对任意固定的$\varepsilon > 0$, 找$\mathscr{A}_2$中的集$O_n \supset A_n$使

$$\mu(O_n) \leq \mu(A_n) + \varepsilon 2^{-n}.$$

再找$\mathscr{A}_1$中的紧集$K \subset A$使

$$\mu(A) \leq \mu(K) + \varepsilon$$

由于

$$K \subset \cup_{n=1}^{\infty} O_n,$$

由有限覆盖定理, 存在N使得

$$K \subset \cup_{n=1}^{N} O_n.$$

因此

$$\mu(K) \le \sum_{n=1}^{N} \mu(O_n)$$

从而

$$\begin{aligned}\mu(A) &\le \sum_{n=1}^{\infty} \mu(O_n) + \varepsilon \\ &\le \sum_{n=1}^{\infty} \mu(A_n) + 2\varepsilon\end{aligned}$$

令$\varepsilon \downarrow 0$就得到要证的不等式.

定理12.5.1的证明:

任取S的一可数基, 其生成的代数记为$\mathscr{A}_0$。则$\mathscr{A}_0$也可数且$\sigma(\mathscr{A}_0) = \mathscr{S}$. 由$P$的正则性,对每一$A \in \mathscr{A}_0$, 存在上升紧集列$\{K_n\}$使

$$P(K_n) \uparrow P(A).$$

这样得到的集合全体$\mathscr{A}_1$依然可数, 记它们与$\mathscr{A}_0$生成的代数为$\mathscr{A}$, 则$\mathscr{A}$也可数. 对$A \in \mathscr{A}$, 任取$P(A|\mathscr{G})$的一等价形$q(\omega, A)$. 于是存在P-零集$N_1 \in \mathscr{G}$, 使对$\omega \notin N_1$, $q(\omega, \cdot)$为$\mathscr{A}$上非负有限可加集函数且$q(\omega, S) = 1$.

若$A \in \mathscr{A}_0$, 则存在$K_n \uparrow$, $K_n \in \mathscr{A}_1$, $A \supset K_n$, 使

$$P(K_n) \uparrow P(A)$$

则

$$1_{K_n} \uparrow 1_A, \quad a.s.$$

由条件期望的单调收敛定理, 有

$$q(\omega, K_n) \uparrow q(\omega, A), \quad a.s.$$

所以存在P-零集N_2使上式对$\omega \notin N_2$及$A \in \mathscr{A}_0$成立. 这样,由引理12.5.4, 对$\omega \notin N_1 \cup N_2$, $q(\omega, \cdot)$为$\mathscr{A}_0$上的测度,因而可扩张到$\mathscr{S}$上.

任取$\mathscr{G}$上一概率测度Q, 定义$(S, \mathscr{S})$上的函数

$$p(\omega, F) = \begin{cases} q(\omega, F), & \omega \notin N_1 \cup N_2 \\ Q(F), & \omega \in N_1 \cup N_2 \end{cases}$$

容易验证p满足定理中的所有要求.

13 概率测度的弱收敛

§13.1 定义及基本性质

在关于积分号下取极限的诸多结果中,我们总是固定测度,而让被积函数变化. 然而应用中往往还需要考虑固定被积函数,而让测度变化的情况. 例如在初等概率论中我们知道,如果一列分布函数弱收敛到一个分布函数的话,那么相应的特征函数就是处处收敛的,反之也对.

现在,代替$\mathbb{R}$, 考虑一般的度量空间上的概率测度的弱收敛问题. 设(X,ρ)度量空间,记其上的Borel代数为$\mathscr{B}$. X上的有界连续函数空间记为$C_b(X)$.

对$A \subset X$, A的内部记为A^o, 闭包记为$\bar{A}$, 边界记为∂A. 设μ是$\mathscr{B}$上测度,如果$\mu(\partial A)=0$, 则称A是μ的连续集.

定义13.1.1. 设$P, P_1, P_2, \cdots$为$(X,\mathscr{B})$上的概率测度. 若

$$\lim_{n\to\infty}\int f dP_n = \int f dP, \quad \forall f \in C_b(X),$$

则称P_n弱收敛于P, 记为$P_n \Rightarrow P$.

由下面的命题, 弱收敛的极限是唯一的.

命题13.1.2. $(X,\mathscr{B})$上的两个测度P与Q若满足

$$\int f dP = \int f dQ, \quad \forall f \in C_b(X),$$

则$P \equiv Q$.

证明: 设A为一闭集, 令

$$f_\varepsilon(x) = g(\frac{1}{\varepsilon}\rho(x,A)), \quad \varepsilon > 0,$$

其中

$$\rho(x,A) = \inf\{\rho(x,y):\ y \in A\},$$

$$g(t) = \begin{cases} 1, & t = 0, \\ 1-t, & t \in [0,1], \\ 0, & t \geq 1. \end{cases}$$

则$f_\varepsilon \in C_b(X)$, $f_\varepsilon \downarrow 1_A$, $\forall \varepsilon > 0$.

在

$$\int f_\varepsilon dP = \int f_\varepsilon dQ$$

中, 令$\varepsilon \downarrow 0$, 由单调收敛定理得到

$$P(A) = Q(A).$$

因此P和Q在所有闭集上相等. 再用单调类定理便完成了证明.

像普通数列的收敛一样,我们有:

命题13.1.3. *若序列$\{P_n\}$的任意子列均包含一个弱收敛到P的子子列,则$P_n \Rightarrow P$.*

证明:设$P_n \not\Rightarrow P$, 则存在$f \in C_b(X)$, $\varepsilon > 0$, 及一子列$n_k \to \infty$, 使得

$$|\int f dP - \int f dP_{n_k}| > \varepsilon, \quad \forall k.$$

故在$\{P_{k_k}\}$中不可能抽出弱收敛到P的子列. 证毕.

下面是弱收敛概念的几种等价形式.

定理13.1.4. *下列说法等价:*

(i) $P_n \Rightarrow P$.

(ii) $\limsup_n P_n(A) \le P(A)$, $\forall$闭集A.

(iii) $\liminf_n P_n(A) \ge P(A)$, $\forall$开集A.

(iv) $\lim_n P_n(A) = P(A)$, $\forall$ P的连续集A.

证明: (ii)$\Leftrightarrow$(iii)是显然的, 因为考虑余集即可.

(i)$\Rightarrow$(ii) 设A闭, f_ε由上面定义. 令

$$A_\varepsilon := \{x : \rho(x) < \varepsilon\}.$$

由于A闭,所以$\varepsilon \downarrow 0$时$A_\varepsilon \downarrow A$. 于是

$$\limsup_n P_n(A) \le \limsup_n \int f_\varepsilon dP_n = \limsup_n \int f_\varepsilon dP \le P(A_\varepsilon) \downarrow P(A), \quad \varepsilon \downarrow 0.$$

(ii)+(iii)$\Rightarrow$(iv) 设A^o为A的内部, $\bar{A}$为A的闭包. 则对于$P(\partial A) = 0$有

$$\limsup_n P_n(A) \le \limsup_n P_n(\bar{A}) \le P(\bar{A}) = P(A),$$

$$\liminf_n P_n(A) \ge \liminf_n P_n(A^o) \ge P(A^o) = P(A).$$

所以

$$\lim_n P_n(A) = P(A).$$

(iv)$\Rightarrow$(i) 设f为X上有界连续函数, $\sup_{x\in X} |f(x)| < M$. 令

$$D := \{t : \; P(f = t) \neq 0\}.$$

则D为可数集. 因此, 对任意$\varepsilon > 0$, 存在$[-M, M]$的分割:

$$-M = t_0 < t_1 < t_2 < \cdots < t_k = M,$$

使$t_i \notin D$且$\max_i |t_i - t_{i-1}| < \varepsilon$.

令

$$B_i := \{x : t_{i-1} \le f(x) < t_i\} = f^{-1}([t_{i-1}, t_i)).$$

由于f连续, 故$f^{-1}((t_i, t_{i+1}))$是开集,于是

$$\partial B_i \subset f^{-1}(\{t_{i-1}\}) \subset f^{-1}(\{t_i\}).$$

所以

$$P(\partial B_i) = 0.$$

从而

$$\lim_n P_n(B_i) = P(B_i), \quad \forall i = 1, \cdots, k.$$

但

$$\begin{aligned}
& |\int f dP_n - \int f dP| \\
\le \; & |\int f dP_n - \sum_{i=1}^k t_t P_n(B_i)| + |\sum_{i=1}^k t_t P_n(B_i) - \sum_{i=1}^k t_t P_(B_i)| \\
& + |\int f dP - \sum_{i=1}^k t_t P_n(B_i)| \\
\le \; & 2\varepsilon + |\sum_{i=1}^k t_t P_n(B_i) - \sum_{i=1}^k t_t P_(B_i)|
\end{aligned}$$

故

$$\limsup_n |\int f dP_n - \int f dP| \le 2\varepsilon.$$

由ε的任意性便得

$$\lim_n |\int f dP_n - \int f dP| = 0.$$

证毕.

往下我们需要X上的上、下半连续函数的概念. 实函数f, 如果满足对$x \in X$及任意序列$x_n \to x$均有

$$f(x) \le \liminf_n f(x_n)$$

或

$$f(x) \ge \limsup_n f(x_n),$$

则分别称为下、上半连续的. 它们又分别等价于对任意$a \in \mathbb{R}$, 集合

$$\{x : f(x) \le a\}$$

或

$$\{x : f(x) \ge a\}$$

是闭集.

一个有界的上半连续函数可由闭集的示性函数的线性组合下降地逼近.事实上,不妨设$0 \le f \le 1$, 令

$$f_n(x) := 2^{-n} \sum_{k=0}^{2^n-1} 1_{f(x) \ge k2^{-n}},$$

则$f_n(x) \downarrow f(x)$, $\forall x \in X$. 同样地, 有界的下半连续函数可由集的示性函数的线性组合上升地地逼近.

现在我们可以叙述更多的等价条件.

定理13.1.5. *下面任何一个条件都和上定理中的条件等价:*

(v) 对任意有界一致连续函数f,

$$\lim_n \int f dP_n = \int f dP.$$

(vi) 对任意有界上半连续函数f,

$$\limsup_n \int f dP_n \le \int f dP.$$

(vii) 对任意有界下半连续函数f,

$$\liminf_n \int f dP_n \ge \int f dP.$$

证明: 显然(i)⇒(v), 而在(i)⇒(ii)所用的函数f_n均是一致连续的, 所以(v)⇒(ii). (vi)自然蕴含(ii), 而用上面所说的逼近很容易证明(ii)⇒(vi). 同样的道理, (iii)⇔(vii). 证毕.

注13.1.6. *在$(S, \mathscr{S})$上的所有概率测度构成的空间中可定义一个与弱收敛等价的度量,且在此度量下该空间为Polish空间. 见[1, Ch.2]*

我们现在转到$X = \mathbb{R}$的特殊情形. 回忆一下, 对分布函数F_n, F, 如果对F的所有连续点x有$F_n(x) \to F(X)$,则称F_n弱收敛于F. 我们在此记也为$F_n \Rightarrow F$.

而我们还知道,分布函数和$\mathbb{B}^1$上的概率测度是一一对应的.

假设F_n对应P_n, F对应P, 那么这两种弱收敛有什么关系呢?

答案是:

定理13.1.7.

$$\{F_n \Rightarrow F\} \Longleftrightarrow \{P_n \Rightarrow P\}.$$

证明: 设$P_n \Rightarrow P$. 若x为F的连续点,则$(-\infty, x)$为P的连续集. 因此

$$\lim_n F_n(x) = \lim_n P_n((-\infty, x)) = P((-\infty, x)) = F(x).$$

反之,设$F_n \Rightarrow F$. 为此,只要证明对任意开集A有

$$\liminf_n P_n(A) \geq P(A).$$

将A表为至多可数个两两不交的开区间的并集:

$$A = \sum_{k=1}^{\infty} I_k, \quad I_k = (a_k, b_k).$$

固定$\varepsilon > 0$. 由于F的不连续点至多是可数个, 故对每一k, 可取均是其连续点的c_k, d_k使$I'_k = (c_k, d_k) \subset (a_k, b_k)$且

$$P(I_k) \leq P(I'_k) + 2^{-k}\varepsilon.$$

于是

$$\lim_n P_n(I'_k) = P(I'_k) \geq P(I_k) - 2^{-k}\varepsilon.$$

这样由Fatou引理就有

$$\begin{aligned}
\liminf_n P_n(A) &= \liminf_n \sum_k P_n(I_k) \\
&\geq \sum_k \liminf_n P_n(I_k) \\
&\geq \sum_k \liminf_n P_n(I'_k) \\
&= \sum_k P(I'_k) \\
&\geq \sum_k (P(I_k) - 2^{-k}\varepsilon) \\
&= P(A) - \varepsilon.
\end{aligned}$$

由ε的任意性便得

$$\liminf_n P_n(A) \geq P(A).$$

证毕.

§13.2 Skorohod定理

设$(\Omega, \mathscr{F}, Q)$为概率空间，X_n, X是其上随机变量，P_n, P分别是其分布. 设$X_n \to X$ a.s.. 如果A是开集, 那么

$$\liminf_n 1_A(X_n) \geq 1_A(X).$$

因此由Fatou引理

$$P(A) \leq \int \liminf_n 1_A(X_n) dQ \leq \liminf_n \int 1_A(X_n) dQ = \liminf_n P_n(A).$$

故$P_n \Rightarrow P$. 所以几乎处处收敛蕴含了其分布的弱收敛.

反过来呢? 反过来是不一定的. 首先, 分布在同一空间上的随机变量完全可以定义在不同的概率空间上,因此几乎处处收敛的便无从谈起. 其次, 即便它们都定义在同一个概率空间上,也非常可能不几乎处处收敛, 这只要看下面的简单例子就可以了: 设X为满足

$$Q(X = 1) = Q(X = -1) = \frac{1}{2}$$

的随机变量,令

$$X_n = (-1)^n X$$

则$P_n \Rightarrow P$, 但X_n几乎处处不收敛.

如果有人争辩说,在上述例子中,可以找出子列几乎处处收敛到X的话,那我们还可以构造出简单的例子使得无法抽出子列收敛(习题).

尽管如此,Skorohod说了,上述命题在某种意义下还是可以反过来的. Skorohod的结果证明很复杂,我们以下面的简单一点的结果为先导.

以Ω表单位区间$[0, 1]$, $\mathscr{F} = \mathscr{B}^1 \cap \Omega$, Q表$\mathscr{F}$上的Lebesgue测度.

定理13.2.1. *对$(X, \mathscr{B})$上的任意概率测度P, 存在定义在$(\Omega, \mathscr{F}, Q)$上的随机变量,它以P为分布.*

注13.2.2. *在本定理及下一定理中, 完备性可以去掉, 见严加安.*

证明: 设$X_0 := \{x_1, x_2, \cdots\}$为$X$的稠密子集, 令

$$B_{1_i} := \{x : \rho(x, x_i) \leq 1\}.$$

则$X = \cup_i B_{1,i}$. 用通常的将一般的并化为不交并的办法,可以找出可数个两两不交的集合A_{1i}, 其中每一个都是某个B_{1i}的子集——于是每个A_{1i}的直径都小于1——,使得

$$X = \sum A_{1i}.$$

记

$$\mathscr{A}_1 := \{A_{1i}\ i = 1, 2, \cdots\}.$$

同样地, 可以找到可数个集合$\{A_{2i}\}$, 其中每个的直径小于2^{-1}, 使

$$X = \sum A_{2i}$$

我们还可以进一步假定每个A_{2i}都是某个A_{1j}的子集,因为这只要把A_{2i}再进行分解

$$A_{2i} = \sum_j A_{2i}A_{1j}$$

就可以做到. 依次下去,对每一k, 可找到可数个集合$\{A_{ki}\}$, 其中每一个A_{ki}的直径均小于2^{-k} 且是某个$A_{k-1,j}$的子集.

相应于这些$\{A_{ki}\}$, 可对$[0,1]$作分解

$$[0,1] = \sum_i I_{ki}$$

使$Q(I_{ki}) = P(A_{ki})$且

$$\{I_{k+1,u} \subset I_{kv}\} \Longleftrightarrow \{A_{k+1,u} \subset A_{kv}\}, \quad \forall k, u, v$$

对每一A_{ki}, 取并固定一个$x_{ki} \in A_{ki} \cap X_0$. 作随机变量

$$X_k(\omega) = x_{ki}, \quad 若\omega \in I_{ki}.$$

则对任意k, 任意ω, 存在j使$\{X_i(\omega), i \geq k\} \subset A_{kj}$. 因而

$$\rho(X_i, X_j) \leq 2^{-k}, \quad \forall i, j \geq k.$$

所以$\{X_k\}$处处有有限的极限X.

往证X的分布为P. 由弱极限的唯一性,这只需证X_k的分布弱收敛到P.

对每一k, 对任意$x \in X$, 存在唯一的i使$x \in A_{ki}$. 将这一映射记为$x \to i_k(x)$, 则有

$$\rho(x, x_{ki_k(x)}) \leq 2^{-k}, \quad \forall x \in X.$$

设f为X上的有界连续函数, 令

$$f_k(x) := \sum_i f(x_{ki}) 1_{A_{ki}}(x) = f(x_{ki_k(x)}).$$

则

$$\lim_k f_k(x) = f(x), \quad \forall x \in X.$$

因此由控制收敛定理

$$\begin{aligned}\int f(x)P_k(dx) &= \int f(X_k)Q(dx)\\ &= \sum_i f(x_{ki})Q(I_{ki})\\ &= \sum_i f(x_{ki})P(A_{ki})\\ &= \int f_k(x)P(dx)\\ &\to \int f(x)P(dx)\end{aligned}$$

证毕.

定理13.2.3. *若$P_n \Rightarrow P$, 则在$(\Omega, \mathscr{F}, Q)$上存在随机变量X_n及X, 它们分别以P_n及P为分布且$\lim_n X_n = X\ \ a.s.$.*

证明: 像上一定理那样构造A_{ki}, 只是附加要求每一A_{ki}均是P-连续集: 这是可以做到的, 因为对任意$x \in X$, r的单调上升函数$P(\{y : \rho(x,y) < r\})$只有可数个间断点, 而但凡r是其连续点, $\{y : \rho(x,y) < r\}$就是P的连续集. 又注意到下面的事实: 两个连续集的交集仍是连续集, 因为我们一般地有$\partial(A \cap B) \subset (\partial A) \cap (\partial B)$.

对每一P_n做同样的事情. 但多做以下的事情: 对每一k, 在排列$[0,1]$的分解I_{ki}及I_{ki}^n时, 按指标i递增的顺序自左至右排列.

这样我们得到$X_{n,k} \to X_n$, $X_{0,k} \to X$, X_n, X的分布分别是P_n及P, 且

$$\rho(X_{n,k}, X_n) < 2^{-k}, \quad \rho(X_{0,k}, X) < 2^{-k}.$$

由于

$$\sum_i (P(A_{ki}) - P_n(A_{ki})) = 0,$$

所以

$$\begin{aligned}\sum_i |Q(I_{ki}) - Q(I_{ki}^n)| &= \sum_i |P(A_{ki}) - P_n(A_{ki})|\\ &= 2\sum_i{}' (P(A_{ki}) - P_n(A_{ki}))\\ &= 2\sum_i (P(A_{ki}) - P_n(A_{ki}))^+\end{aligned}$$

其中$\sum'$表示对非负项求和. 注意A_{ki}是P的连续集, 因此由控制收敛定理得

$$\lim_n \sum_i |Q(I_{ki}) - Q(I_{ki}^n)| = 0.$$

于是更有对任意u

$$\lim_n \sum_{i<u} |Q(I_{ki}) - Q^n(I_{ki})| = 0.$$

固定k, u, 设α, α_n分别是I_{ku}及I_{ku}^n的左端点, 则

$$\alpha = \sum_{i<u} Q(I_{ki}) = \lim_n \sum_{i<u} Q(I_{ki}^n) = \lim_n \alpha_n.$$

因此若$\omega \in I_{ki}^o$, 则当n充分大时, $\omega \in (I_{ki}^n)^o$. 于是有

$$\rho(X(\omega), X_n(\omega)) \le 2^{1-k}.$$

所以若ω不是任何一个I_{ki}的端点, 必有

$$\lim_n X_n(\omega) = X(\omega).$$

其余所有I_{ki}的端点是一可数集,因而其Lebesgue测度为零. 这样就完成了证明.

Skorohod的这个定理有许多用处. 我们来看几个它的直接推论:

推论13.2.4. *若X_n, X是取值于$(X, \mathscr{B})$的随机变量(可以定义在不同的概率空间上), $X_n \Rightarrow X$, 则存在概率空间$(\Omega, \mathscr{F}, P)$及定义在其上的,分别与X_n, X同分布的随机变量Y_n, Y使得$\lim_n Y_n = Y$ a.s..*

再设$(S_1, \mathscr{S}_1)$与$(S_2, \mathscr{S}_2)$是两个完备可分度量空间, P_n, P是$(S_1, \mathscr{S}_1)$上的概率测度. 设$f: S_1 \mapsto S_2$为$\mathscr{S}_1/\mathscr{S}_2$可测. 用$D_f$表示$f$的不连续点.

推论13.2.5. *若$P_n \Rightarrow P$且$P(D_f) = 0$, 则$P_n f^{-1} \Rightarrow P f^{-1}$.*

推论13.2.6. *设$X_n \Rightarrow X$且$P(X \in D_f) = 0$, 则$f(X_n) \Rightarrow f(X)$.*

推论13.2.7. *设$X_n \Rightarrow X$, 则*

$$E[|X|] \le \liminf_n E[|X_n|].$$

推论13.2.8. *设$X_n \Rightarrow X$且X_n一致可积, 则X一致可积且*

$$\lim_n E[X_n] = E[X].$$

§13.3 Prohorov定理

设S为Polish空间, $\mathscr{S}$是S上的Borel代数. 考虑$(S, \mathscr{S})$上的概率测度. 我们下面一个问题是: 什么样的概率测度序列有弱收敛的子列? 或者说,在弱收敛的意义下是相对紧的(以下简称相对紧)?

回答这个问题的是Prohorov. 为此我们先引进胎紧性的概念.

定义13.3.1. $(S,\mathscr{S})$上的概率测度P称为胎紧的, 若对任意$\varepsilon>0$, 存在紧集K, 使得

$$P(K)>1-\varepsilon.$$

回忆一个集合如果能写成一列紧集的并集,则称为σ紧. 这样由定义直接得到,当且仅当一个测度被一个σ紧集所支撑时,它是胎紧的.

因此,如果S本身是σ紧的(例如$\mathbb{R}^n$就是如此), 则其上的任意概率测度都是胎紧的. 一般地,我们有下面的

定理13.3.2. $\{P_n\}$相对紧的充要条件是: 它是胎紧的.

这个定理现在有很简洁的证明. 但为了避免使用一些拓扑学和泛函分析的深刻结果,我们这里选用一个更初等然而也更直接,更测度论的证明.

定理的证明:胎紧$\Longleftrightarrow$相对紧.

(I)设$S=\mathbb{R}^1$. 由于概率测度序列的弱收敛等价于相应的分布函数的弱收敛加上极限函数是分布函数, 故首先由Helly第一定理得到有子序列弱收敛,然后又由胎紧性得到极限函数确实是分布函数.

(II)设$S=\mathbb{R}^n$. 此时的证明与$\mathbb{R}^1$几乎一样,留作练习.

(III) $S=\mathbb{R}^\infty$. 这里

$$\mathbb{R}^\infty:=\{(x_1,x_2,\cdots),\quad x_i\in\mathbb{R}\},$$
$$\rho(x,y)=\sum_k 2^{-k}\frac{|x_k-y_k|}{1+|x_k-y_k|}.$$

在这个度量下$\mathbb{R}^\infty$显然是完备可分的，且

$$\{x^{(n)}\to x\}\Longleftrightarrow\{x_i^{(n)}\to x_i,\ \forall i\}.$$

定义

$$\pi_k:\mathbb{R}^\infty\to\mathbb{R}^k:\quad \pi((x_1,x_2,\cdots))=(x_1,x_2,\cdots,x_k).$$

则π_k是连续映射, 故由推论13.2.5, $(\mathbb{R}^k,\mathscr{B}^k)$上的概率测度族$\{P_n\pi_k^{-1}\}$是胎紧的, 故存在$(\mathbb{R}^k,\mathscr{B}^k)$上的概率测度$Q_k$及子列$\{k_n\}$使

$$P_{k_n}\pi_k^{-1}\to Q_k.$$

由对角线方法, 可抽出公共的子列$\{n'\}$使

$$P_{n'}\pi_k^{-1}\to Q_k,\ \forall k.$$

显然$\{Q_k\}$满足Kolmogorov相容性定理的条件, 因而在$(S,\mathscr{S})$上存在概率测度Q使得

$$Q\pi_k=Q_k,\quad \forall k.$$

于是

$$P_{n'}\pi_k^{-1} \to Q\pi_k,\ \forall k.$$

现在证明

$$P_{n'} \to Q.$$

设$\{x_1^0, x_2^0, \cdots\}$为$\mathbb{R}^\infty$的稠密子集$O \subset \mathbb{R}^\infty$为开集, 则$O$可表示为

$$O = \cup_{i=1}^\infty A_i$$

其中每个A_i都具有形状

$$A_i = \pi_i^{-1}(B_i)$$

而$B_i \subset \mathbb{R}^i$可表示为

$$B_i = \{(x_1, \cdots, x_i) : |x_k - x_k^0| < \varepsilon_i, \quad \forall k \le i\}.$$

其中$\varepsilon_i > 0$且B_i为Q_i的连续集. 这样

$$P_{n'}\pi_i^{-1}(B_i) \to Q_i(B_i).$$

设$\varepsilon > 0$. 选m使得

$$Q(O) \le Q(\cup_{i=1}^m A_i) + \varepsilon$$

由于$\pi_m(\cup_{i=1}^m A_i)$是Q_m的连续集, 故有

$$\begin{aligned} Q(O) - \varepsilon &\le Q(\cup_{i=1}^m A_i) \\ &\le Q_m(\pi_m \cup_{i=1}^m A_i) \\ &= \lim_{n'} P_{n'}(\cup_{i=1}^m A_i) \\ &\le \liminf_{n'} P_{n'}(\cup_{i=1}^\infty A_i) \\ &= \liminf_{n'} P_{n'}(O). \end{aligned}$$

令$\varepsilon \downarrow 0$便完成了证明.

(IV)对一般的S, 我们需要下面的拓扑结果.

引理13.3.3. *S同胚于$\mathbb{R}^\infty$的某个Borel集.*

实际上可证明更强的结果: S同胚于$[0,1]^\infty$的某个G_δ集. 见Rogers and Williams.

设P_n胎紧，令

$$Q_n = P_n h^{-1}.$$

由于h连续，故S中紧集在h下的像为$\mathbb{R}^\infty$中的紧集，于是Q_n胎紧. 因此存在Q使得(不妨设)$Q_n \Rightarrow Q$. 又对任意$\varepsilon > 0$, 存在S中的紧集K使得

$$P_n(K) \ge 1 - \varepsilon$$

所以$\mathbb{R}^\infty$中的紧集$h(K)$使得

$$Q_n(h(K)) \geq 1-\varepsilon$$

故

$$Q(h(K)) \geq 1-\varepsilon.$$

所以

$$Q(h(S)) = 1.$$

于是令

$$P = Qh$$

则P为S上的概率测度且

$$P_n \Rightarrow P.$$

(用推论13.2.5). 至此定理证毕.

注13.3.4. 证明中没有用到完备性. 而用所谓σ紧集过渡, 也可去掉可分性的限制, 见*Billingsley.*

§13.4 空间C

§13.5 空间D

14 独立性

在初等概率论里面我们已经接触过两个或一般地有限个随机变量间的独立性的概念, 这个概念可以归纳进下面所讲的σ代数间的独立性中.

设概率空间为$(\Omega, \mathscr{F}, P)$.

定义14.0.1. 设$\mathscr{G}_1$, $\mathscr{G}_2$是$\mathscr{F}$的子σ代数, 若

$$g_1 \in L^2(\mathscr{G}_1),\ g_2 \in L^2(\mathscr{G}_2),\ E[g_1] = E[g_2] = 0 \Longrightarrow\ E[g_1 g_2] = 0, \qquad (14.0.28)$$

则称$\mathscr{G}_1$与$\mathscr{G}_2$是独立的.

设ξ_1, ξ_2是两个随机变量. 若$\sigma(\xi_1)$, $\sigma(\xi_2)$是独立的, 则称ξ_1, ξ_2是独立的.

命题14.0.2. 下面五个条件都是$\mathscr{G}_1$, $\mathscr{G}_2$独立的充要条件是

i)

$$g_1 \in L^2(\mathscr{G}_1),\ g_2 \in L^2(\mathscr{G}_2) \Longrightarrow\ E[g_1 g_2] = E[g_1]E[g_2];$$

ii)

$$g_1 \in L^\infty(\mathscr{G}_1),\ g_2 \in L^\infty(\mathscr{G}_2) \Longrightarrow\ E[g_1 g_2] = E[g_1]E[g_2];$$

iii) $\forall A_1 \in \mathscr{G}_1,\ A_2 \in \mathscr{G}_2$, 有

$$P(A_1A_2) = P(A_1)P(A_2).$$

iv)

$$g \in L^1(\mathscr{G}_1) \Longrightarrow\ E[g|\mathscr{G}_2] = E[g].$$

v)

$$g_1 \in L^1(\mathscr{G}_1),\ g_2 \in L^1(\mathscr{G}_2) \Longrightarrow\ g_1g_2 \in L^1(\mathscr{F})\ 且 E[g_1g_2] = E[g_1]E[g_2];$$

证明:

$i) \Longrightarrow$ 14.0.28是明显的; 14.0.28 $\Longrightarrow i)$只要用$g_i - E[g_i]$代替g_i即可.

$i) \Longrightarrow ii)$也是明显的; $ii) \Longrightarrow i)$用老一套, 即用一个函数的截尾逼近它自己.

$ii) \Longrightarrow iii)$ 还是明显的; $iii) \Longrightarrow ii)$用控制收敛定理.

v)自然蕴涵ii); 为证明ii)蕴涵v), 假定g_1, g_2都是非负的(一般情况可分为正部与负部讨论). 此时有$g_{1,n}, g_{2,n} \in L^\infty$使$g_{1,n} \uparrow g_1$, $g_{2,n} \uparrow g_2$. 由2)知等式先对$g_{1,n}, g_{2,n}$成立, 再由单调收敛定理得到对g_1, g_2成立.

最后, 用定理12.2.2, 有$iii) \Longrightarrow iv)$. 完毕.

这个定理的一个直接推论是, 这里给出的随机变量的独立性的定义与本科教材中的定义是一致的.

对于多个的情形, 独立性的定义是:

定义14.0.3. *设$\mathscr{G}_i, 1 \le i \le n$为$\mathscr{F}$的$n$个子$\sigma$代数, $H \in \{1, \cdots, n\}$, 令*

$$\mathscr{G}_H := \sigma\{\mathscr{G}_i, i \in H\}.$$

若$\forall H$, $\mathscr{G}_H$与$\mathscr{G}_{H^c}$独立, 则称$\mathscr{G}_i, 1 \le i \le n$相互独立.

又设I是任一指标集, $\mathscr{G}_i, i \in I$为$\mathscr{F}$的n个子σ代数, 若对I的任意有限子集K, $\mathscr{G}_i, i \in K$独立, 则称$\mathscr{G}_i, i \in I$独立.

从命题14.0.2出发, 用归纳法不难证明:

命题14.0.4. *下面三条件都是$\mathscr{G}_1, \cdots, \mathscr{G}_n$独立的充要条件是*

i)

$$g_i \in L^1(\mathscr{G}_i), \Longrightarrow\ \prod_{i=1}^n g_i \in L^1(\mathscr{F})\ 且 E[\prod_{i=1}^n g_i] = \prod_{i=1}^n E[g_i];$$

ii)

$$g_i \in L^\infty(\mathscr{G}_i) \Longrightarrow E[\prod_{i=1}^n g_i] = \prod_{i=1}^n E[g_i];$$

iii) $\forall A_i \in \mathscr{G}_i$, 有

$$P(\cap_{i=1}^n A_i) = \prod_{i=1}^n P(A_i).$$

15 鞅

§15.1 定义及基本性质

设$(\Omega, \mathscr{F}, P)$是概率空间, $\xi_i, i = 1, 2, \cdots$是其上的中心(即$E[\xi_i] = 0$)独立随机变量序列, 令

$$S_n := \sum_{i=1}^{n} \xi_i,$$

$$\mathscr{F}_n := \sigma\{\xi_i, 1 \le i \le n\},$$

则有

$$E[S_n|\mathscr{F}_{n-1}] = E[S_{n-1} + \xi_n|\mathscr{F}_n] = E[S_{n-1}|\mathscr{F}_{n-1}] + E[\xi|\mathscr{F}_{n-1}] = S_{n-1}.$$

从这个不等式出发, 我们就得到鞅的概念.

定理15.1.1. *设$\mathscr{F}_n$是$\mathscr{F}$的一列上升的子σ代数序列, $\forall n$, $M_n \in L^1(\mathscr{F})$且$M_n \in \mathscr{F}_n$. 若*

$$E[M_{n+1}|\mathscr{F}_n] = M_n,$$

则$\{M_n, n \ge 1\}$称为鞅. 若$\forall n$, $M_n \in L^p$, 则称为局部L^p鞅; 若还有$\sup_n \|M_n\| < \infty$, 则称为L^p鞅.

命题15.1.2. *设$\{M_n, n \ge 1\}$称为鞅, 则$\forall m > n$, 有*

$$E[M_m|\mathscr{F}_n] = M_n.$$

证明: 设$m = n + p$, 则

$$E[M_{n+p}|\mathscr{F}_n] = E[E[M_{n+p}|\mathscr{F}_{n+p-1}]|\mathscr{F}_n] = E[M_{n+p-1}|\mathscr{F}_n] = \cdots = E[M_{n+1}|\mathscr{F}_n] = M_n.$$

完毕.

§15.2 L^2鞅

由于L^2为Hilbert空间, 因此L^2鞅处理起来是比较方便的. 首先我们有下面的"能量等式"

命题15.2.1. *设$\{M_n, \mathscr{F}_n\}$为局部L^2鞅, 则$\forall m > n$有*

$$E[M_m^2] - E[M_n^2] = \sum_{i=n}^{m} E[(M_{i+1} - M_i)^2].$$

证明: 只需在$m=n+1$时证明. 但此时由于

$$E[M_nM_{n+1}]=E[E[M_nM_{n+1}|\mathscr{F}_n]]=E[M_nE[M_{n+1}|\mathscr{F}_n]]=E[M_n^2],$$

故

$$\text{右边}=E[M_n^2-2M_nM_{n+1}+M_{n+1}^2]=E[M_{n+1}^2]-E[M_n^2]=\text{左边}.$$

完毕.

由于上式右边非负, 故有

推论15.2.2. *设$\{M_n,\mathscr{F}_n\}$为局部L^2鞅, 则$E[M_n^2]$是单调上升的序列; 它还是L^2鞅, 则$\lim_{n\to\infty}E[M_n^2]$存在.*

又由于将一个鞅抽掉若干项后, 它仍然是鞅, 故又有

推论15.2.3. *设$\{M_n,\mathscr{F}_n\}$为局部L^2鞅, 则$\forall m>n$有*

$$E[M_m^2]-E[M_n^2]=E[(M_m-M_n)^2].$$

记

$$\mathscr{F}_\infty:=\sigma\{\mathscr{F}_n,n\geq 1\},$$

关于L^2鞅我们有下面重要的

定理15.2.4. *(结构定理) 设$\{M_n,\mathscr{F}_n\}$是L^2鞅, 则存在$M_\infty\in L^2(\mathscr{F}_\infty)$使*

(i) $\|M_n-M_\infty\|_2\to 0$;

(ii) $M_n=E[M_\infty|\mathscr{F}_n]$.

(iii)$\sum_{i=1}^\infty(M_{i+1}-M_i)^2<\infty$ a.s..

反之, 设$\mathscr{G}_n$是任意上升的$\mathscr{F}$的子σ代数序列, $\mathscr{G}_\infty:=\sigma\{\mathscr{G}_n,n\geq 1\}$, $f\in L^2(\mathscr{G}_\infty)$. 令

$$N_n:=E[f|\mathscr{G}_n].$$

则$\{N_n,\mathscr{G}_n\}$是L^2鞅且

$$\|N_n-f\|\to 0.$$

证明: 因为$E[M_n^2]$单调上升而又有上界, 故有有限的极限, 因而是Cauchy列. 再由

$$\|M_n-M_m\|_2^2=|E[M_m^2]-E[M_n^2]|,$$

知M_n是$L^2(\mathscr{F})$中的Cauchy列. 由于$L^2(\mathscr{F})$是完备的, 故有$M_\infty\in L^2(\mathscr{F})$使在$L^2(\mathscr{F})$中$M_n\to M_\infty$. 又因为存在$M_n$的子列几乎处处收敛到$M_\infty$, , 故$M_\infty\in\mathscr{F}_\infty$, 因而$M_\infty\in L^2(\mathscr{F}_\infty)$. (i)得证.

又根据鞅性和条件期望的连续性, 有

$$M_n = E[M_{n+k}|\mathscr{F}_n] = \lim_{k\to\infty} E[M_{n+k}|\mathscr{F}_n] == E[\lim_{k\to\infty} M_{n+k}|\mathscr{F}_n] = E[M_\infty|\mathscr{F}_n].$$

(ii)得证.

至于(iii), 由命题15.2.1

$$\sum_{i=1}^{\infty} E[(M_{i+1}-M_i)^2] = \sup_n E[M_n^2] - E[M_1^2] < \infty,$$

所以(iii)成立.

反之, 由于$\forall n$,有

$$E[N_{n+1}|\mathscr{G}_n] = E[E[f|\mathscr{G}_{n+1}]|\mathscr{G}_n] = E[f|\mathscr{G}_n] = N_n,$$

所以N_n是鞅. 又

$$E[N_n^2] = E[(E[f|\mathscr{G}_n])^2] \le E[E[f^2|\mathscr{G}_n]] = E[f^2],$$

所以N_n是L^2鞅. 于是由已证明的第一部分知存在$N_\infty \in L^2(\mathscr{G})$使在$L^2(\mathscr{G}_\infty)$中$N_n \to N_\infty$且$N_n = E[N_\infty|\mathscr{G}_n]$. 下面证明$N_\infty = f$.

我们只需证明: $\forall G \in \mathscr{G}_\infty$有

$$\int_G N_\infty \, dP = \int_G f \, dP.$$

令$\mathcal{E}$是使此等式成立的$\mathscr{G}_\infty$中的元素全体, 则$\mathcal{E}$为单调类. 而$\forall n$, $\forall G \in \mathscr{G}_n$, 有

$$\int_G N_\infty dP = \int_G N_n dP = \int_G f dP,$$

所以$\mathcal{E} \supset \cup_{n=1}^\infty \mathscr{G}_n$. 由于$\mathscr{G}_n$是上升的, 故$\cup_{n=1}^\infty \mathscr{G}_n$是代数. 这样由单调类定理便有$\mathcal{E} = \mathscr{G}_\infty$. 完毕.

§15.3 L^1鞅

比起L^2鞅的理论, L^1鞅要复杂得多. 为研究他们我们需要停时的概念.

定理15.3.1. 设$\{\mathscr{F}_n\}$是$\mathscr{F}$的递增子σ代数列, T是$Omega$上取正整数值的函数. 若$\forall n$, 有

$$\{\omega, T(\omega) > n\} =: A_{T,n} \in \mathscr{F}_n,$$

或, 等价地,

$$\{\omega, T(\omega) = n\} =: A_{T,n} \in \mathscr{F}_n,$$

则称T为停时.

所谓停时, 它的直观解释是: 若把T看成某个事件发生的时刻, 则此时刻现在是否到来, 凭到现在为止掌握的信息即可做出判断.

例如, 一个人第一次结婚的时刻是停时, 但最后一次结婚的时刻就不是停时.

命题15.3.2. 设T_1, T_2是停时, 则

$$T_3 := T_1 \wedge T_2$$

也为停时.

证明:

$$\{T_3 > n\} = \{T_1 > n\} \cap \{T_2 > n\}.$$

完毕.

定理15.3.3. 设$\mathcal{M}_n$是鞅, T是停时. 令

$$M_n^T(\omega) := M_{T(\omega)\wedge n}(\omega).$$

则M_n^T为鞅, 称为截尾鞅.

证明:我们有

$$M_n^T = \sum_{i=1}^{n-1}(M_{i+1} - M_i)1_{A_{T,i}} + M_1.$$

上式右边的每个函数都关于$\mathscr{F}_n$可测, 因而左边关于$\mathscr{F}_n$可测. 而

$$\begin{aligned} E[M_{n+1}^T|\mathscr{F}_n] &= \sum_{i=1}^{n-1}(M_{i+1} - M_i)1_{A_{T,i}} + M_1 + E[(M_{n+1} - M_n)1_{A_{T,n}}|\mathscr{F}_n] \\ &= \sum_{i=1}^{n-1}(M_{i+1} - M_i)1_{A_{T,i}} + M_1 + E[(M_{n+1} - M_n)|\mathscr{F}_n]1_{A_{T,n}} \\ &= \sum_{i=1}^{n-1}(M_{i+1} - M_i)1_{A_{T,i}} + M_1 = M_n^T. \end{aligned}$$

因此M_n^T是鞅. 完毕.

利用Doob停止定理可以得到也是属于Doob的鞅的极大不等式. 我们先引进记号:

$$M_n^* := \sup_{1\le i\le n} |M_i|,$$

$$M^* := \sup_{i\ge 1} |M_i| = \lim_{n\to\infty} M_n^*.$$

我们有

定理15.3.4. 设M_n为鞅, 则$\forall\gamma > 0$, 有

$$(i) \qquad P(M_n^* \geq \gamma) \leq \frac{2}{\gamma}[E(|M_n|) + E(M_1|)],\ \forall n$$

及

$$(ii) \qquad P(M^* \geq \gamma) \leq \frac{4}{\gamma}\sup_n E(|M_n|).$$

证明: 令

$$A_\gamma^n := \{\omega : \sup_{i<n} M_i(\omega) \geq \gamma\},$$

$$T(\omega) = \begin{cases} \inf\{i : M_i(\omega) \geq \gamma\} & \text{if } \omega \in A_\gamma^n \\ n & \text{if } \omega \notin A_\gamma^n \end{cases}$$

则$T \leq n$且

$$\{\omega : T(\omega) > j\} = \cup_{i \leq j}\{\omega : M_i(\omega) < \gamma\} \in A_j \text{ 若} j < n,$$

而

$$\{\omega : T(\omega) > n\} = \emptyset.$$

因而T是停时. 以M_n^T记其T-截尾鞅, 则由于$M_1^T = M_1$, 故

$$E[M_n^T] = E[M_1].$$

而

$$E[M_n^T] = E[M_n^T 1_{\{T<n\}}] + E[M_n^T 1_{\{T=n\}}].$$

由于M_n^T在$\{T < n\}$上大于等于γ, 在$\{T = n\}$上等于Y_n, 所以

$$E[M_1^T] - E[M_n 1_{\{T=n\}}] = E[M_n^T 1_{\{T<n\}}] \geq \gamma E[1_{\{T<n\}}] = \gamma P(A\gamma^n).$$

于是更有

$$\gamma P(A\gamma^n) \leq E[|M_1^T|] + E[|M_n|].$$

考虑$-M_n$同样可得

$$\gamma P(\sup_{i<n}[-M_i] \geq \gamma) \leq E[|M_1^T|] + E[|M_n|].$$

上面两式相加就得到(i). 再注意$\forall\varepsilon > 0$,

$$\{M^* \geq \gamma\} \subset \lim_{n\to\infty}\{M_n^* \geq \gamma - \varepsilon\},$$

所以

$$P(M^* \geq \gamma) \leq \lim_{n\to\infty} P(M_n^* \geq \gamma - \varepsilon) \leq \frac{4}{\gamma - \varepsilon}\sup_n E[|M_n|],$$

由ε的任意性得到(ii). 完毕.

下面讨论L^1鞅的收敛性. 先给出

定义15.3.5. 设$\{M_n, \mathscr{F}_n\}$为鞅. 若存在$M_\infty \in L^1(\mathscr{F}_\infty)$使

$$M_n = E[f|\mathscr{F}_n],$$

则称$\{M_n, \mathscr{F}_n\}$为正则鞅, M_∞称为M_n的终值.

对于正则鞅我们有如下的

定理15.3.6. *(收敛定理)*设$\{M_n, \mathscr{F}_n\}$为正则鞅, 终值为M_∞. 则

$$\lim_{n\to\infty} \|M_n - M_\infty\|_1 = 0.$$

且

$$\lim_{n\to\infty} M_n = M_\infty \ a.s..$$

证明: 取$N_m \in L^2(\mathscr{F}_\infty)$使$\lim_{m\to\infty} \|N_m - M_\infty\|_1 = 0$. 令

$$M_n^m = E[N_m|\mathscr{F}_n].$$

则$\forall m$, $\{\mathcal{M}_n^m, n \geq 1\}$为$L^2$鞅且

$$\lim_{n\to\infty} \|M_n^m - N_m\|_2 = 0.$$

但我们有

$$\begin{aligned}
\|M_n - M_\infty\|_1 &\leq \|M_n - M_n^m\|_1 + \|M_n^m - N_m\|_1 + \|N_m - M_\infty\|_1 \\
&\leq \|M_\infty - N_m\|_1 + \|M_n^m - N_m\|_1 + \|N_m - M_\infty\|_1 \\
&\leq 2\|M_\infty - N_m\|_1 + \|M_n^m - N_m\|_2.
\end{aligned}$$

先令$n \to \infty$再令$m \to \infty$即得

$$\lim_{n\to\infty} \|M_n - M_\infty\|_1 = 0.$$

再证几乎处处收敛性. $\forall i$, 令

$$\beta_i(\omega) := \sup_{n>i} |M_n(\omega) - M_i(\omega)|,$$

则只需证明$\lim_{i\to\infty} \beta_i = 0$. 令

$$\beta(\omega) := \lim_{i\to\infty} \beta(\omega).$$

由于

$$X_n^i := M_{n+i} - M_n$$

关于$\mathscr{F}_{n+i}$为鞅, 故$\forall\gamma > 0$,

$$P(\beta_i > \gamma) \le \frac{4}{\gamma}\sup_n \|X_n^i\|_1.$$

但

$$\lim_{i\to\infty}\sup_n \|X_n^i\|_1 = \lim_{i\to\infty}\sup_n \|M_{n+i} - M_n\|_1 = 0,$$

所以

$$\lim_{i\to\infty} P(\beta_i > \gamma) = 0.$$

由于β_q下降到β, 故

$$P(\beta > \gamma) = \lim_{i\to\infty} P(\beta_i > \gamma) = 0.$$

因而$\mathcal{M}_n$几乎处处收敛. 由于在L^1中$M_n \to M_\infty$, 所以$\lim_{n\to\infty} M_n = M_\infty$. 完毕.

现在研究L^1鞅的收敛性. 我们先引进停止σ代数的概念.

定义15.3.7. 设T是停时, 令

$$\mathscr{F}_T := \{E \in \mathscr{F}_\infty : E \cap \{T \le n\} \in \mathscr{F}_n \quad \forall n\},$$

或者等价地

$$\mathscr{F}_T := \{E \in \mathscr{F}_\infty : E \cap \{T = n\} \in \mathscr{F}_n \quad \forall n\},$$

称为停止于T的σ代数.

读者自己容易看出这两个条件的确是等价的. 这里的直观解释是: $\mathscr{F}_T$包含到T时刻为止的所有信息. 例如, 若T表示第一次生病住院的时刻, 则$\mathscr{F}_T$的意思是: 在n时刻或之前住院的人的直到n时刻为止的所有情况.

命题15.3.8. 设$\forall n, \quad X_n \in \mathscr{F}_n$, T为停时, 则$X_T \in \mathscr{F}_T$.

证明: $\forall B \in \mathbb{R}$,

$$\{X_T \in B\} = \cup_{n=1}^{\infty}(\{T = n\} \cap \{X_n \in B\} \in \mathscr{F}_T.$$

完毕.

命题15.3.9. 设σ, τ为两停时且$\sigma \le \tau$, 则$\mathscr{F}_\sigma \subset \mathscr{F}_\tau$.

证明:$\forall E \in \mathscr{F}_\sigma$, $\forall n$, 我们有

$$E \cap \{\tau = n\} = E \cap \{\sigma < n\} \cap \{\tau = n\} \in \mathscr{F}_n.$$

完毕.

定理15.3.10. *(Doob停止定理)若$\{M_n, n \geq 1\}$为鞅, σ, τ为有界停时且$\sigma \leq \tau$, 则*

$$M_\sigma = E[X_\tau | \mathscr{F}_\sigma].$$

证明: 先设$\sigma \leq \tau \leq N$, $\tau - \sigma \leq 1$. 则对$E \in \mathscr{F}_\sigma$, 则$\forall k = 1, \cdots, n-1$, 有

$$E \cap \{\sigma = k\} \cap \{\tau = k+1\} = E \cap \{\sigma = k\} \cap \{\tau > k\} \in \mathscr{F}_k.$$

所以

$$\begin{aligned}
\int_E M_T \, dP &= \sum_{k=1}^n \int_{E \cap \{\sigma=k\} \cap \{\tau=k\}} M_T \, dP + \sum_{k=1}^n \int_{E \cap \{\sigma=k\} \cap \{\tau=k+1\}} M_T \, dP \\
&= \sum_{k=1}^n \int_{E \cap \{\sigma=k\} \cap \{\tau=k\}} M_k \, dP + \sum_{k=1}^n \int_{E \cap \{\sigma=k\} \cap \{\tau=k+1\}} M_{k+1} \, dP \\
&= \sum_{k=1}^n \int_{E \cap \{\sigma=k\} \cap \{\tau=k\}} M_k \, dP + \sum_{k=1}^n \int_{E \cap \{\sigma=k\} \cap \{\tau=k+1\}} M_k \, dP \\
&= \sum_{k=1}^n \int_{E \cap \{\sigma=k\}} M_k \, dP \\
&= \int_E M_S \, dP.
\end{aligned}$$

一般地, 令

$$\tau_k = \tau \vee (\sigma + k), \ k = 0, 1, \cdots, n-1.$$

则τ_k均是停时, $\sigma = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_n = \tau$,且$\tau_k - \tau_{k-1} \leq 1$. 于是

$$E[M_\tau | \mathscr{F}_\sigma] = E[E[M_\tau | \mathscr{F}_{\tau_{n-1}}] | \mathscr{F}_s] = E[M_{\tau_{n-1}} | \mathscr{F}_\sigma] = \cdots = M_\sigma.$$

完毕.

下面我们要考虑一般的(未必是L^2)鞅收敛性问题. 这里要用到下面的思想:

设x_n是有界一数列, a, b是两实数且$a < b$. 记

$$n_1 = \min\{n : x_n \leq a\};$$

$$n_2 = \min\{n \geq n_1 : x_n \geq b\};$$

....

$$n_{2k+1} = \min\{n \geq n_{2k} : x_n \leq a\};$$

$$n_{2k+2} = \min\{n \geq n_{2k+1} : x_n \geq b\};$$

这里我们约定$\min \emptyset = \infty$. 这样在n_1到n_2之间, x_n从下到上穿越了$[a, b]$一次; 在n_2到n_3之间, x_n从上到下穿越了$[a, b]$一次. 若x_n收敛, 则$\forall a < b$, 必有k_0使$n_{k_0} = \infty$, 即x_n只能穿越$[a, b]$有限次; 反之, 若x_n不收敛, 则存在两个子列y_n, z_n分别收敛到不同的极限, 设他们分别是a与b. 不妨设$a < b$. 于是x_n必须穿越区间$[a + \frac{1}{4}(b-a), b - \frac{1}{4}(b-a)]$无穷次. 于是我们得到

命题15.3.11. 设x_n是有界一数列, 则x_n收敛的充要条件是x_n穿越任何有限区间的次数都是有限的; 这又等价于, x_n穿越任何以有理数为端点的有限区间的次数都是有限的.

我们将用这命题来研究鞅的收敛性. 对鞅$\{M_n\}$, 令

$$\tau_0 = 0;$$

$$\tau_1 = \min\{n : M_n \leq a\};$$

$$\tau_2 = \min\{n \geq \tau_1 : M_n \geq b\};$$

$$\cdots\cdots\cdots$$

$$\tau_{2k+1} = \min\{n \geq \tau_{2k} : M_n \leq a\};$$

$$n_{2k+2} = \min\{n \geq \tau_{2k+1} : M_n \geq b\};$$

$$\cdots\cdots\cdots$$

则诸τ_k为停时. 再令

$$U_m(a, b) := \max\{k : \tau_{2k} \leq m\}.$$

则$U_m(a, b)$为随机变量.

命题15.3.12. 我们有

$$E[U_m(a, b)] \leq (b-a)^{-1} E[(M_m - a)^+ - (X_0 - a)^+].$$

为证明此命题, 我们需要下鞅的概念.

定义15.3.13. 设$\{X_n, n \geq 1\}$是随机变量序列, $X_n \in L^1(\mathscr{F}_n)$, $\forall n$. 若

$$E[X_{n+1}|\mathscr{F}_n] \geq X_n, \quad \forall n,$$

则称X_n为下鞅.

鞅当然是下鞅了. 注意从平均的角度来看, 下鞅是上升的而非下降的, 因此下鞅这个名称似乎有点怪怪的. 但大家现在都习惯这样叫了. 习惯的势力总是强大的.

由关于条件期望的Jensen不等式, 有

命题15.3.14. 设$\{X_n, n \geq 1\}$为下鞅. 若ϕ是凸函数且$\forall n$, $\phi(X_n)$可积, 则$\{\phi(X_n), n \geq 1\}$为下鞅.

注意对凸函数$\phi(x) = |x|$, x^+, 上述命题是永远可以用的; 但对于诸如$\phi(x) = |x|^p$, $p > 1$之类的凸函数, 使用此命题前就必须先验证可积性.

和关于鞅的停止定理15.3.10一样, 我们有下面的结果, 其证明就略去了.

定理15.3.15. 若$\{X_n, n \geq 1\}$为下鞅, σ, τ为有界停时且$\sigma \leq \tau$, 则

$$X_\sigma \leq E[X_\tau | \mathscr{F}_\sigma].$$

命题15.3.12的证明:

证明: 令$X_n = (M_n - a)^+$, 则M_n上穿(a, b)的次数等于X_n上穿$(0, b-a)$的次数. 即

$$U_m(X, 0, b-a) = U_m(M, a, b).$$

再令$\tau_n' := \tau_n \vee m$, 则$2k > m$时有

$$\begin{aligned} X_m - X_0 &= \sum_{n=1}^{2k} (X_{\tau_n'} - X_{\tau_{n-1}'}) \\ &= \sum_{n=1}^{k} (X_{\tau_{2n}'} - X_{\tau_{2n-1}'}) + \sum_{n=0}^{k-1} (X_{\tau_{2n+1}'} - X_{\tau_{2n}'}) \end{aligned}$$

所以

$$E[X_m] - E[X_0] = E[\sum_{n=1}^{k} (X_{\tau_{2n}'} - X_{\tau_{2n-1}'})] + \sum_{n=0}^{k-1} E[(X_{\tau_{2n+1}'} - X_{\tau_{2n}'})]$$

上式中自然有$\sum_{n=1}^{k} (X_{\tau_{2n}'} - X_{\tau_{2n-1}'}) \geq (b-a) U_m(X, b-a) = (b-a) U_m(M, a, b)$. 要小心的是第二个期望: 虽然$X_{\tau_{2n+1}} - X_{\tau_{2n}} \leq a - b < 0$, 但未必有$X_{\tau_{2n+1}'} - X_{\tau_{2n}'} < 0$. 相反, 由命题15.3.14 及定理15.3.15, 有

$$E[X_{\tau_{2n+1}'} - X_{\tau_{2n}'}] \geq 0.$$

这样我们就得到

$$E[X_m] - E[X_0] \geq (b-a) E[U_m(M, a, b)],$$

命题得证. 完毕.

现在我们可以证明

定理15.3.16. 若$\{M_n, n \geq 1\}$为L^1鞅, 则存在$M_\infty \in L^1(\mathscr{F}_\infty)$使$M_\infty = \lim_{n\to\infty} M_n$ *a.s.*.

证明: 以Q表示全体有理数的集合. 令

$$\begin{aligned} E: &= \{\omega: \exists r, s \in Q,\ r < s,\ 使U_\infty(r,s)(\omega) = \infty\} \\ &= \cup_{r,s\in Q}\{\omega : U_\infty(r,s)(\omega) = \infty\}. \end{aligned}$$

我们先证$P(E) = 0$. 为此只需要证明$\forall r, s \in \mathbb{R}$, $\forall n$, 有

$$E[U_\infty(r,s)] < \infty \qquad\qquad (*).$$

但由刚刚证明的命题, 对任意m有

$$E[U_m(r,s)] < (r-s)^{-1} \sup_n \|M_n\|_1.$$

令$m \to \infty$, 由单调收敛定理就得到(*), 因而$P(E) = 0$, 于是$M_\infty = \lim_{n\to\infty} M_n$几乎处处存在. 再由Fatou引理便知$M_\infty \in L^1(\mathscr{F}_\infty)$. 完毕.

要注意的是, 尽管此定理断言了极限的几乎处处存在性和可积性, 却并没有M_n是正则鞅即说$M_n = E[M_\infty|\mathscr{F}_n]$. 事实上, 这一般是不成立的: 人们已经构造了这样的L^1鞅, 它非零, 然而几乎处处收敛到零.

要使$M_n = E[M_\infty|\mathscr{F}_n]$, 必须付出代价.

定理15.3.17. *设$\{M_n, n \geq 1\}$为L^1鞅, $M_\infty = \lim_{n\to\infty} M_n$, a.s.. 则下列条件等价:*

(i) $\{M_n, n \geq 1\}$为正则鞅;

(ii)M_n一致可积;

(iii)$\|M_n - M_\infty\|_1 = 0$.

为证明它, 我们下面的准备工作:

定义15.3.18. *若随机变量族$\{f_i, i \in I\}$满足*

$$\lim_{c\to\infty} \sup \int_{|f_i|\geq c} |f_i|\ dP = 0,$$

则称为一致可积族.

一致可积族的概念和我们以前学的积分一致绝对连续有密切的关系. 这就是

命题15.3.19. *随机变量族$\{f_i, i \in I\}$一致可积的充要条件是: 其积分有界且一致绝对连续.*

证明: 必要性. 设$\{f_i, i \in I\}$一致可积, 于是存在$M > 0$使

$$\sup_i \int_{|f_i|\geq M} |f_i|\ dP \leq 1.$$

因此

$$\sup_i \int |f_i|\ dP \le 1 + M,$$

即积分是有界的. 又$\forall \varepsilon > 0$, 有$c > 0$使

$$\sup_i \int_{|f_i| \ge c} |f_i|\ dP \le \frac{\varepsilon}{.}$$

再取$\delta := \varepsilon c^{-1}$, 则$P(F) \le \delta$时有

$$\begin{aligned}
\sup_i \int_F |f_i|\ dP &= \sup_i \int_{F \cap \{|f_i| \le c\}} |f_i|\ dP + \sup_i \int_{F \cap \{|f_i| > c\}} |f_i|\ dP \\
&\le cP(F) + \sup_i \int_{\{|f_i| > c\}} |f_i|\ dP \\
&\le 2\varepsilon.
\end{aligned}$$

积分的一致绝对连续性得证.

充分性. 设积分有界且一致绝对连续, 则$\forall \varepsilon > 0$, 存在$\delta > 0$使$P(F) \le \delta$时

$$\sup_i \int_F |f_i|\ dP \le \varepsilon.$$

由于$c > \delta^{-1} \sup_i E[|f_i|]$时有(由Chebushev不等式)

$$\sup_i P(|f_i| > c) \le \delta,$$

因而此时

$$\sup_i \int_{\{|f_i| > c\}} |f_i|\ dP \le \varepsilon.$$

故一致可积. 完毕.

引理15.3.20. *设$f \in L^1$, I是任意指标集, $\{\mathscr{F}_i, i \in I\}$是$\mathscr{F}$的子$\sigma$代数族. 令$f_i = E[f|\mathscr{F}_i]$, 则$\{f_i, i \in I\}$是一致可积的.*

我们需要一个预备结果.

引理15.3.20的证明: $\forall \varepsilon > 0$, 由积分的绝对连续性, 知存在$\delta > 0$使$P(A) < \delta$时有

$$\int_A |f|\ dP < \varepsilon.$$

取$a = \delta^{-1} E[|f|]$, 则由Chebyshev不等式有

$$P(|f| > a) < \delta.$$

因而
$$\int_{\{|f|>a\}} |f|\ dP < \varepsilon.$$
再取$N_\varepsilon = \varepsilon^{-1} a E[|f|]$, 则当$N > N_\varepsilon$时就有
$$\begin{aligned}\int_{\{|f_i|>N\}} |f_i|\ dP &\leq \int_{\{|f_i|>N\}} |f|\ dP \\ &\leq aP(|f_i| > N) + \int_{\{|f|>a\}} |f|\ dP \\ &\leq aN^{-1}E[|f_i|] + \varepsilon \\ &= \varepsilon + \varepsilon = 2\varepsilon.\end{aligned}$$
完毕.

定理的证明: (i)$\Longrightarrow$(ii): 由正则鞅的定义, 有$M_n = E[M_\infty|\mathscr{F}_n]$. 于是由引理$\{M_n, n \geq 1\}$一致可积.

(ii)$\Longrightarrow$(iii): 由积分的收敛定理.

(iii)$\Longrightarrow$(i): $\forall n$有
$$M_n = \lim_{m\to\infty} E[M_m|\mathscr{F}_n] = E[M_\infty|\mathscr{F}_n].$$
完毕.

§15.4 应用: Radon-Nikodym定理

设f是$[0,1]$上的绝对连续函数. 由实变函数论, f'关于Lebesgue测度几乎处处存在且
$$f(x) = \int_0^x f'(t)dt.$$
设μ_f是f产生的Lebesgue-Stieltje测度. 从上式出发, 用单调类定理容易证明, 对任意Lebesgue可测集A均有
$$\mu_f(A) = \int_A f\ dt.$$
下面我们要将这个结果推广到一般的概率空间.

首先需要推广的是绝对连续的概念.

定义15.4.1. *设$(\Omega, \mathscr{F}, P)$是概率空间, Q是$\mathscr{F}$上的有限测度. 若*
$$P(A) = 0 \Longrightarrow Q(A) = 0, \quad \forall A \in \mathscr{F},$$
则称Q关于P绝对连续. 记为$Q \leq P$.

绝对连续性的这种定义表面上看来似乎和函数的绝对连续性的定义不太一样. 但事实上我们有

命题15.4.2. $Q \le P$的充要条件是: $\forall \varepsilon > 0$, $\exists \delta > 0$, 使

$$P(A) < \delta \Longrightarrow Q(A) < \varepsilon. \qquad (*)$$

证明: 充分性是显然的, 只需证必要性.

若$(*)$不成立, 则存在$\varepsilon_0 > 0$, 使$\forall k > 0$, $\exists A_k \in \mathscr{F}$满足

$$P(A_k) < 2^{-k}$$

而

$$Q(A_k) \ge \varepsilon_0.$$

于是, 令

$$A := \limsup_k A_k = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k.$$

则有

$$P(A) = \lim_{n\to\infty} P(\cup_{k=n}^{\infty} A_k) \le \lim_{n\to\infty} \sum_{k=n}^{\infty} 2^{-k} = 0,$$

而

$$Q(A) = \lim_{n\to\infty} Q(\cup_{k=n}^{\infty} A_k) \ge \varepsilon_0.$$

因此Q不关于P绝对连续.

本小节中, 我们将假定$(\Omega, \mathscr{F}, P)$是可分概率空间. 所谓可分性是指:

定义15.4.3. 设$(\Omega, \mathscr{F}, P)$是概率空间. 若$L^1(\Omega, \mathscr{F}, P)$是可分的(即$L^1$有一个稠密子集$\{f_n, n \ge 1\}$), 则称$(\Omega, \mathscr{F}, P)$是可分的.

当然不是所有的概率空间都是可分的, 因此可分性实实在在是一个限制条件. 但这个条件又是如此之弱: 在通常的应用中它总是满足的.

我们有:

命题15.4.4. $(\Omega, \mathscr{F}, P)$是可分概率空间的充要条件是: 存在$\mathscr{F}$的上升的有限σ子代数列$\{\mathscr{F}_n, n \ge 1\}$使对任意$f \in L^1$有

$$f = \lim_{n\to\infty} E[f|\mathscr{F}_n].a.s., \qquad (*)$$

或等价地

$$f = \lim_{n\to\infty} E[f|\mathscr{F}_n] \quad (L^1). \qquad (**)$$

此时称$\{\mathscr{F}_n, n \ge 1\}$ P-生成$\mathscr{F}$.

证明: 首先注意由于$\{E[f|\mathscr{F}_n], n \geq 1\}$是一致可积鞅, 因此$(*)$与$(**)$确实是等价的.

先证明必要性. 设$(\Omega, \mathscr{F}, P)$是可分概率空间, 则存在L^1的稠密子集$\{f_n, n \geq 1\}$. 再用简单函数逼近每个f_n, 知存在可数个简单函数$\{g_n, n \geq 1\}$在L^1中稠密.

令

$$\mathscr{F}_n := \sigma\{g_i,\ 1 \leq i \leq n\}.$$

则$\{\mathscr{F}_n, n \geq 1\}$为上升的有限$\sigma$代数列. 由正则鞅的收敛定理, $\forall m$有

$$g_m = \lim_{n\to\infty} E[g_m|\mathscr{F}_n].$$

设$f \in L^1$. $\forall \varepsilon > 0$, 取m使

$$\|f - g_m\|_1 < \varepsilon.$$

再取N使$n > N$时有

$$\|g_m - E[g_m|\mathscr{F}_n]\|_1 < \varepsilon.$$

则$n > N$时,

$$\begin{aligned} \|f - E[f|\mathscr{F}_n]\|_1 &< \|f - g_m\|_1 + \|g_m - E[g_m|\mathscr{F}_n]\|_1 + \|E[g_m|\mathscr{F}_n] - E[f|\mathscr{F}_n]\|_1 \\ &\leq 2\|f - g_m\|_1 + \|g_m - E[g_m|\mathscr{F}_n]\|_1 < 3\varepsilon. \end{aligned}$$

因此

$$f = \lim_{n\to\infty} E[f|\mathscr{F}_n].$$

反之, 设对任意的$f \in L^1$上式满足, 则f可由$\cup_n L^1(\mathscr{F}_n)$中的函数逼近, 而后者是可数的. 完毕.

现在可以叙述本小节的主要结果了.

定理15.4.5. *(Radon-Nikodym定理) 设$(\Omega, \mathscr{F}, P)$是可分概率空间, Q是$\mathscr{F}$上的有限测度. 则下列两条件等价:*

(i) $Q \leq P$;

(ii)存在$k \in L^1$, $k \geq 0$, 使

$$Q(A) = \int_A k\ dP.$$

上述的k称为Q对P的Radon-Nikodym导数, 记为$k = \frac{dQ}{dP}$.

由积分的性质, (ii)是蕴涵了(i)的. 因此我们要证明的是(i)$\Longrightarrow$(ii). 让我们从最简单的情况开始.

引理15.4.6. 设$\mathscr{F}$是有限σ代数, 其原子为$\{A_i, 1 \le i \le n\}$. 又设$Q \le P$. 则

$$k(\omega) = \frac{Q(A_i)}{P(A_i)},$$

这里我们约定$\frac{0}{0} = 0$.

这是可以直接验证的, 读者不妨自己试试.

引理15.4.7. 设$\{\mathscr{F}_n, n \ge 1\}$是$\mathscr{F}$的上升的有限子$\sigma$代数族, P_n、Q_n分别表示P、Q在$\mathscr{F}_n$上的限制. 令

$$k_n := \frac{dQ_n}{dP_n},$$

则$\{k_n, n \ge 1\}$是一致可积鞅.

证明: 任意固定n. 设$\mathscr{F}_n$的原子为$\{A_1, \cdots, A_p\}$, $\mathscr{F}_{n+1}$的原子为$\{B_1, \cdots, B_q\}$. 自然有$p \le q$.

设$\omega \in A_i$. 由于$A_i \in \mathscr{F}_{n+1}$, 故它可以写成并集:

$$A_i = \cup_{i=1}^{j} B_{n_i}, \quad \{n_1, \cdots, n_j\} \subset \{1, \cdots, q\}.$$

因此

$$E[k_{n+1}|\mathscr{F}_n](\omega) = \frac{\sum_{i=1}^{j} \int_{B_{n_i}} k_{n+1}\ dP}{P(A_i)} = \frac{\sum_{i=1}^{j} Q(B_{n_i})}{P(A_i)} = \frac{Q(A_i)}{P(A_i)} = k_n(\omega).$$

这样便证明了$\{k_n, n \ge 1\}$是鞅. 下面证明它是一致可积的.

首先, 由于它是非负的, 故是L^1鞅, 因此只需要证明其积分是一致绝对连续的.

$\forall E \in \mathscr{F}$, $\forall n$, 设$\mathscr{F}_n$的原子为$A_1, \cdots, A_l$, 则

$$\int_A |k_n|\ dP = \sum_{i=1}^{l} \frac{Q(A_i)}{P(A_i)} P(A_i \cap A) = \sum_{i=1}^{l} \frac{P(A_i \cap A)}{P(A_i)} Q(A_i);$$

$$P(A) = \sum_{i=1}^{l} \frac{P(A_i \cap A)}{P(A_i)} P(A_i).$$

令

$$\mathcal{A}_s := \{A_i : \frac{P(A_i \cap A)}{P(A_i)} \in [2^{-s-1}, 2^{-}s).\},$$

$$B_s := \cup_{A_i \in \mathcal{A}_s} A_i.$$

则

$$\int_A |k_n|\ dP \le \sum_{s=1}^{\infty} 2^{-s} Q(B_s);$$

$$P(A) \geq \sum_{s=1}^{\infty} 2^{-s-1} P(B_s).$$

现在, $\forall \varepsilon > 0$, 取s_0使$2^{-s_0+1}Q(\Omega) < \frac{\varepsilon}{2}$, 则

$$\int_A |k_n| \; dP \leq \sum_{s=1}^{s_0} 2^{-s} Q(B_s) + \frac{\varepsilon}{2};$$

由于$Q \leq P$, 存在δ'使$P(A) < \delta'$时$Q(A) < \frac{\varepsilon}{4}$. 再取$\delta = 2^{-s_0-1}\delta$, 则$P(A) < \delta$时, 对$0 \leq s < s_0$就有

$$P(B_s) \leq 2^{s_0+1} P(A) < \delta'.$$

因而

$$Q(B_s) < \frac{\varepsilon}{4}.$$

这样

$$\int_A |k_n| \; dP \leq \sum_{s=0}^{s_0} 2^{-s} \frac{\varepsilon}{4} + \frac{\varepsilon}{2} < \varepsilon.$$

注意上述δ只依赖于ε而不依赖于n, 因此就证明了积分的一致绝对连续性. 完毕.

现在我们可以给出

Radon-Nikodym定理的证明: 由上一引理, k_n是一致可积鞅, 因而是正则鞅. 设其终值为k. 我们来证明

$$\int_A k \; dP = Q(A), \quad \forall A \in \mathscr{F}.$$

由单调类定理容易证明, 对于任意$\phi \in L^\infty(\mathscr{F}_n)$, 有

$$\int k_n \phi \; dP = \int \phi \; dQ.$$

取$\phi_n = E[1_A|\mathscr{F}_n]$就得到

$$\int k_n E[1_A|\mathscr{F}_n] \; dP = \int \phi_n \; dQ,$$

即

$$\int_A k_n \; dP = \int \phi_n \; dQ.$$

$n \to \infty$时, 由于$\|k - k_n\|_1 \to 0$, 故左边趋于$\int_A k \; dP$; 另一方面,由于ϕ_n是P-a.e.趋于1_A的, 故由$Q \leq P$也是Q-a.e.趋于1_A的; 再加上ϕ_n永远不超过1, 因而由控制收敛定理右边趋于$\int_A \; dQ = Q(A)$. 因而

$$\int_A k \; dP = Q(A).$$

完毕.

注: 若P只是有限测度而未必是概率测度时, 定理照样成立. 这是因为$P(\Omega)^{-1}P$必然是概率测度, 而我们只要考虑此测度即可. 同样地, 下面的结果也不限于概率测度.

命题15.4.8. 设$Q \le P$, 则对任意非负可测函数f有

$$\int f \ dQ = \int f \frac{dQ}{dP} dP.$$

证明只是简单地由定义出发, 用一下单调类定理.

对Radon-Nikodym导数, 可以像对普通的导数一样进行"约分".

命题15.4.9. 设P_1, P_2, P_3为有限测度且$P_1 \le P_2 \le P_3$,则

$$\frac{dP_3}{dP_1} = \frac{dP_3}{dP_2} dP_2 dP_1.$$

证明: $\forall A \in \mathscr{F}$, 我们有

$$P_3(A) = \int_A \frac{dP_3}{dP_2} dP_2 = \int_A \frac{dP_3}{dP_2} \frac{dP_2}{dP_1} dP_1.$$

完毕.

注: 其实对未必可分的测度空间, Radon-Nikodym定理也是成立的. 例如见Halmos等.

§15.5 下鞅的Doob分解

§15.6 L^p对偶

设$p > 1$, 我们来找出L^p的对偶空间.

定理15.6.1. 设$(X, \mathscr{F}, \mu)$为σ有限测度空间, $p > 1$, 则$L^p(X, \mathscr{F}, \mu)$的对偶可等同于$L^q(X, \mathscr{F}, \mu)$, 其中q是p的共轭指数, 即$p^{-1} + q^{-1} = 1$.

证明: $\forall f \in L^q$, 令

$$l_f(g) := \int_X fg \ d\mu.$$

则l_f是定义在L^p上的线性泛函且由Hölder不等式有

$$|l_f(g)| \le \|f\|_q \|g\|_p. \qquad (*)$$

因此l_f是L^p上的连续线性泛函, 即$l_f \in (L^p)^*$. 下面要证明映射$f \mapsto l_f$既是满射又是单射. 这又等价于: 它是保范的满射.

保范性是容易的: 由(*)有

$$\|l_f\|_{(L^p)^*} \le \|f\|_q.$$

然而若取

$$g := \mathrm{sgn}(f)|f|^{p/q}\|f\|_p^{p/q},$$

则有

$$\|g\|_p = 1$$

且

$$l_f(g) = \|f\|_q.$$

所以

$$\|l_f\|_{(L^p)^*} = \|f\|_q.$$

剩下我们还需要证明$f \mapsto l_f$为满射. 即要证明:

(*)*任给*$l \in (L^p)^*$, *存在*$f \in L^q$, *使得*$l = l_f$

我们给出两种证明方法: 一种是利用我们刚刚学过的Radon-Nikodym定理, 另一种则是不用. 之所以要这样做, 一方面是大家可以比较一下, 看看方法的不一样会怎样简化问题的解决; 其次是那个复杂的方法里包含了一个重要的概念, 既L^p的一致凸性, 以及由此而产生的一些令人愉快的结果.

方法一. 首先注意到, 可设μ为有限测度, 因为否则的话,可利用μ的穷竭序列, 在每一个测度有限的子空间上考虑.

先设l是L^p上的正泛函, 即

$$l(f) \geq 0, \qquad \forall f \in L^p, \quad f \geq 0.$$

$\forall A \in \mathscr{F}$, 由于$1_A \in L^p$, 故可以令

$$\nu(A) := l(1_A).$$

由l的线性性和连续性和正性易证ν是$\mathscr{F}$上的有限测度. 又当$\mu(A) = 0$时, 1_A作为L^p的元素为0, 故$\nu(A) = 0$. 因此$\nu \leq \mu$. 于是存在$f \in L^1$使

$$\nu(A) = \int_A f \ d\mu.$$

我们来证明事实上有

$$f \in L^q.$$

令

$$f_n := f 1_{|f| \leq n},$$

$$u_n := \|f_n\|_q^{p^{-1}q} f_n^{p^{-1}q}.$$

则$\|u_n\|_p = 1$. 于是$l(u_n) \le \|l\|$. 但令一方面

$$l(u_n) = \int u_n f \ d\mu = \int u_n f_n \ d\mu = \|f_n\|_q.$$

故由Fatou引理

$$\int f^q \ d\mu = \int \lim_{n\to\infty} f_n^q \ d\mu \le \lim_{n\to\infty} \int f_n^q \ d\mu = |l(u_n)|^q \le \|l\|^q < \infty,$$

故$f \in L^q$.

由定义, l_f与l在示性函数上相等, 于是在简单函数上相等. 由于全体简单函数在L^p中稠密, 故$l_f \equiv l$.

方法二. 我们先建立:

命题15.6.2. *(Clarkson不等式)设$p > 1$, f, $g \in L_p$.*

(i) 若$p \ge 2$, 则

$$\|\frac{f+g}{2}\|_p^p + \|\frac{f-g}{2}\|_p^p \le \frac{1}{2}(\|f\|_p^p + \|g\|_p^p);$$

(ii)若$1 < p \le 2$, 则

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f-g}{2}\|_p^q \le [\frac{1}{2}(\|f\|_p^p + \|g\|_p^p)]^{\frac{1}{p-1}}.$$

其中q是p的共轭指数.

为证明此命题, 需要一下引理:

引理15.6.3. *设$x \in [0,1]$.*

(i) 若$p \ge 2$, 则

$$(\frac{1+x}{2})^p + (\frac{1-x}{2})^p \le \frac{1}{2}(1+x^p);$$

(ii) 若$1 < p \le 2$, 则

$$(1+x)^q + (1-x)^q \le 2(1+x^p)^{\frac{1}{p-1}}.$$

其中q是p的共轭指数.

证明: 先看(i). 令

$$F(x) := (\frac{1+x}{2})^p + (\frac{1-x}{2})^p - \frac{1}{2}(1+x^p).$$

要证明在$[0,1]$上有$F \le 0$. 注意

$$F(0) = \frac{1}{2}(2^{-p+2} - 1) \le 0.$$

故只需证明在$(0,1]$上有$F \leq 0$. 这又只需证明

$$\phi(x) := \frac{2}{x^p} F(x) \leq 0, \quad x \in (0,1].$$

直接计算知$Phi(1) = 0$, 因此只需证明Phi在$(0,1)$上单调上升, 而这只要$\forall x \in (0,1)$, $Phi'(x) \geq 0$就行了.

计算得

$$\phi'(x) = -\frac{p}{x^{p+1}}[(1+x)^{p-1} + (1-x)^{p-1} - 2^{p-1}].$$

因此只要证

$$\psi(x) := (1+x)^{p-1} + (1-x)^{p-1} - 2^{p-1} \leq 0, \quad x \in (0,1).$$

但由于$\psi(1) = 0$, 而

$$\psi'(x) := (p-1)(1+x)^{p-2} + (p-1)(1-x)^{p-2} \geq 0, \quad x \in (0,1),$$

所以确实有$\psi(x) \leq 0$, $\forall x \in (0,1)$. (i)证毕.

(ii)的证明要复杂一些, 但也只涉及到一些初等的计算.

首先, $p = 2$时显然是成立的, 因此下面假设$1 < p < 2$.

$x = 0, 1$时, 不等式是显然的. 对$x \in (0,1)$, 做变换

$$x = \frac{1-u}{1+u},$$

则u跑遍$(0,1)$时, x也跑遍$(0,1)$, 因此(ii)等价于

$$(1 + \frac{1-u}{1+u})^q + (1 - \frac{1-u}{1+u})^q \leq 2(1 + (\frac{1-u}{1+u})^p)^{\frac{1}{p-1}}, \quad u \in (0,1).$$

这又等价于(上式两边乘$(1+u)^q$):

$$2^q(1+u^q) \leq 2[(1+i)^p + (1-u)^p]^{\frac{1}{p-1}}, \quad u \in (0,1).$$

两边取$p-1$次幂就成为

$$(1+u^q)^{p-1} \leq \frac{1}{2}[(1+u)^p + (1-u)^p], \quad u \in (0,1).$$

用幂级数展开的方法可以证明, 这最后一个不等式确实是成立的, 但证明相当烦琐, 我们略去它. 有兴趣的读者可参考E. Hewitt与K. Stromberg的Real and Abstract Analysis.

我们还需要

命题15.6.4. *($0<p<1$时的Minkowski不等式)* 设$f, g \geq 0$, $0<p<1$, 则

$$\|f+g\|_p \geq \|f\|_p + \|g\|_p.$$

它的证明与$p \geq 1$时是类似的.

现在我们可以给出Clarkson不等式的证明:

(i) $p \geq 2$时, 由引理15.6.3对任何实数x, y有

$$|\frac{x+y}{2}|^p + |\frac{x-y}{2}|^p \leq \frac{1}{2}(|x|^p + |y|^p).$$

因而

$$|\frac{f(x)+g(x)}{2}|^p + |\frac{f(x)-g(x)}{2}|^p \leq \frac{1}{2}(|f(x)|^p + |g(x)|^p).$$

两边积分即得

$$\|\frac{f+g}{2}\|_p^p + \|\frac{f-g}{2}\|_p^p \leq \frac{1}{2}(\|f\|_p^p + \|g\|_p^p).$$

(ii)若$1<p\leq 2$, 则由$0<p<1$时的Minkowski不等式有

$$\||\frac{f+g}{2}|^q\|_{p-1} + \||\frac{f+g}{2}|^q\|_{p-1} \leq \||\frac{f+g}{2}|q + |\frac{f-g}{2}|q\|_{p-1}.$$

即

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f+g}{2}\|_p^q \leq \||\frac{f+g}{2}|^q + |\frac{f-g}{2}|^q\|_{p-1}.$$

但上式右边等于

$$\left[\int \left(|\frac{f+g}{2}|^q + |\frac{f-g}{2}|^q\right)^{p-1} \ d\mu\right]^{\frac{1}{p-1}}.$$

它小于等于

$$\left[\int 2^{p-1}(|\frac{f}{2}|^p + |\frac{g}{2}|^p \ d\mu\right]^{\frac{1}{p-1}} = \left[\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\right]^{\frac{1}{p-1}}.$$

完毕.

我们知道, 紧集上的连续函数一定有极大值, 但一般非集上的连续函数却未必有极大值. 因此, 对一般的Banach空间B, 其上的连续线性泛函的范数, 作为B中单位球上的值的绝对值的上确界, 是不一定能够取到的. 然而具体到L_p, 此上确界恰恰可以取到.

定理15.6.5. 设$l \in (L^p)^*$, 则存在$g \in L^p$, $\|g\|_p = 1$, 使$l(g) = \|l\|$.

证明: 如$\|l\|=0$, 则随便取哪个g都行, 因此下面假定$\|l\|>0$.

由定义, 存在L^p中的序列g'_n使

$$\|g'_n\|_p=1,\quad \lim_{n\to\infty}|l(g'_n)|=\|l\|,\quad |l(g'_n)|>\frac{1}{2}\|l\|.$$

为了去掉绝对值, 令$g_n=\mathrm{sgn}(l(g'_n))g'_n$, 则有

$$\|g_n\|_p=1,\quad \lim_{n\to\infty}l(g_n)=\|l\|,\quad l(g'_n)>\frac{1}{2}\|l\|.$$

我们来证明$\{g_n,n\geq 1\}$为L^p中的Cauchy列.

若否, 则有正数$\varepsilon>0$及两子列$\{g_{n_k}\}$及$\{g_{m_k}\}$使

$$\|g_{n_k}-g_{m_k}\|_p>\varepsilon,\quad \forall k.$$

由Clarkson不等式有, $p\geq 2$时,

$$\|\frac{g_{n_k}+g_{m+k}}{2}\|_p^p+\|\frac{g_{n_k}-g_{m+k}}{2}\|_p^p\leq\frac{1}{2}\|g_{n_k}\|_p^p+\frac{1}{2}\|g_{m_k}\|_p^p=1;$$

$1<p<2$时

$$\|\frac{g_{n_k}+g_{m+k}}{2}\|_p^q+\|\frac{g_{n_k}-g_{m+k}}{2}\|_p^q\leq\left[\frac{1}{2}\|g_{n_k}\|_p^p+\frac{1}{2}\|g_{m_k}\|_p^p\right]^{\frac{1}{p-1}}=1.$$

他们分别意味着, 在$p\geq 2$时

$$\|\frac{g_{n_k}+g_{m+k}}{2}\|_p^p\leq 1-\left(\frac{\varepsilon}{2}\right)^p;$$

$1<p<2$时

$$\|\frac{g_{n_k}+g_{m+k}}{2}\|_p^q\leq 1-\left(\frac{\varepsilon}{2}\right)^q.$$

因此, 对任意$p>1$, 总能找到一个与k无关的常数ε'使

$$\|\frac{g_{n_k}+g_{m+k}}{2}\|_p^p\leq 1-\varepsilon',\quad \forall k.$$

由于$l(g_n)>0,\quad \forall n$, 故$g_{n_k}+g_{m_k}\neq 0$. 于是可令

$$h_k:=\frac{g_{n_k}+g_{m_k}}{\|g_{n_k}+g_{m_k}\|_p}.$$

我们有

$$\begin{aligned}
l(h_k) &= \frac{1}{\|g_{n_k}+g_{m_k}\|_p}[l(g_{n_k})+l(g_{m_k})]\\
&= \frac{1}{\frac{\|g_{n_k}+g_{m_k}\|_p}{2}}[\frac{1}{2}l(g_{n_k})+\frac{1}{2}l(g_{m_k})]\\
&\geq \frac{1}{1-\varepsilon'}[\frac{1}{2}l(g_{n_k})+\frac{1}{2}l(g_{m_k})].
\end{aligned}$$

令$k \to \infty$得

$$\liminf_{k\to\infty} \geq \frac{1}{1-\varepsilon'}.$$

由于$\|h_k\|_p = 1$, 上式是决不可能得. 矛盾说明了g_n是Cauchy列. 设其极限为g, 则

$$l(g) = \lim_{n\to\infty} l(g_n) = \|l\|.$$

完毕.

引理15.6.6. 设$l \in (L^p)^*$, $g \in L^p$使$\|g\|_p = 1$, $l(g) = \|l\| > 0$. 任给$f \in L^p$, 令

$$\phi_f(t) = \|g + tf\|_p, \quad t \in \mathbb{R}.$$

则当ϕ_f在$t = 0$处可微时, 有

$$\frac{l(f)}{\|l\|} = \phi'(0).$$

证明: 不妨设$\|l\| = 1$. $\forall t$, 有

$$l(g + t(f - l(f)g)) = l(g) + t(l(f) - l(f)) = l(g) = 1.$$

因此

$$\|g + t(f - l(f)g)\|_p \geq 1.$$

由于$|t|$充分小时有

$$g + tf = (1 + tl(f))[g + t\frac{1}{1+tl(f)}(f - l(f)g)],$$

故

$$\|g + tf\|_p - \|g\|_p \geq |1 + tl(f)| - 1 = tl(f).$$

这样

$$\lim_{t\downarrow 0} \frac{\phi_f(t) - \phi_f(0)}{t} \geq l(f),$$

$$\lim_{t\uparrow 0} \frac{\phi_f(t) - \phi_f(0)}{t} \leq l(f).$$

所以

$$\lim_{t\to 0} \frac{\phi_f(t) - \phi_f(0)}{t} = l(f).$$

完毕.

现在我们回到(*)得证明. 设u, v是两实数, $p > 1$. 考虑$[-1, 1]$上得函数

$$\alpha(t) := |tu + v|^p.$$

我们有

$$\alpha'(t) = pu(tu+v)|tu+v|^{p-2}.$$

因此由中值定理有

$$||tu+v|^p - |u|^p| \leq (|u|+|v|)^p |t|.$$

现在令

$$\beta(t) = \int |tf+g|^p \ d\mu,$$

则有

$$\frac{\beta(t)-\beta(0)}{t} = \int \frac{|tf+g|^p - |g|^p}{t} \ d\mu.$$

由于

$$|\frac{|tf+g|^p - |g|^p}{t}| \leq (|f|+|g|)^p,$$

故由控制收敛定理有

$$\lim_{t\to 0} \frac{\beta(t)-\beta(0)}{t} = \int pfg|g|^{p-2}.$$

于是

$$\begin{aligned} \phi_f'(0) &= \frac{1}{p}\left(\int |g|^p \ d\mu\right)^{\frac{1}{p}-1} \beta'(0) \\ &= \int fg|g|^{p-2} \ d\mu. \end{aligned}$$

因此

$$l(f) = \|l\| \int fg|g|^{p-2} \ d\mu.$$

故取$h := \|l\| g|g|^{p-2}$即有$h \in L^q$且$l = l_h$. 完毕.

§15.7 Lebesgue分解

作为Radon-Nikodym定理的应用, 我们来证明Lebesgue分解定理.

简单起见, 我们设μ、ν均是有限测度. 与绝对连续相关的另一个重要概念是相互奇异性.

定义15.7.1. *若存在$B \in \mathscr{F}$使$\mu(B) = 0$而$\nu(B^c) = 0$, 则称μ与ν互相奇异. 记为$\mu \perp \nu$.*

所以互相奇异这个概念就象军事上的划江而治, 互不干涉. 这里还有两点值得注意: 首先互相奇异是对称的概念, 即双方地位平等, 这与绝对连续是不同的; 其次, 互相奇异和绝对连续并不是互斥的, 比如恒为零的测度, 它既与任何测度互相奇异, 又关于任意测度绝对连续.

任意拿两个测度, 一般它们的关系当然处于中间状态: 既不是互相奇异, 又不是一个关于另一个绝对连续. 不过, *Lebesgue*却发现可以将这种复杂的关系分解为简单关系的和.

定理15.7.2. *(Lebesgue分解定理) 设μ、ν是有限测度, 则存在唯一的分解*

$$\nu = \nu_1 + \nu_2,$$

其中ν_1与ν_2均是有限测度且$\nu_1 \leq \mu$, $\nu_2 \perp \mu$.

证明: 显然$\nu \leq \mu + \nu$. 令

$$f = \frac{d\nu}{d(\mu+\nu)}.$$

对任意$\varepsilon > 1$, 令

$$E_\varepsilon := \{x \in X : f(x) \geq \varepsilon\}.$$

则

$$\nu(E_\varepsilon) = \int_{E_\varepsilon} f \ d(\mu+\nu) \geq \varepsilon\mu(E_\varepsilon) + \varepsilon\nu(E_\varepsilon) \geq \varepsilon\nu(E_\varepsilon).$$

因此

$$\nu(E_\varepsilon) = \mu(E_\varepsilon) = 0.$$

所以

$$(\mu+\nu)(E_\varepsilon) = 0.$$

这样

$$(\mu+\nu)(f > 1) = (\mu+\nu)(\cup_{n=1}^\infty E_{\frac{1}{n}}) \leq \sum_{n=1}^\infty (\mu+\nu)(E_{\frac{1}{n}}) = 0.$$

因此不妨设$f(x) \leq 1$, $\forall x \in X$. 再令

$$B := \{x \in X : \ f(x) = 1\}.$$

由于

$$\nu(B) = \int_B f \ d(\mu+\nu) = \mu(B) + \nu(B),$$

故$\mu(B) = 0$. 令

$$\nu_1(E) := \nu(E \cap B^c), \quad \nu_2(E) := \nu(E \cap B), \quad \forall E \in \mathscr{F}.$$

则$\nu_2 \perp \mu$且

$$\nu = \nu_1 + \nu_2.$$

下面证明$\nu_1 \le \mu$. 设$E \in \mathscr{F}$且$\mu(E) = 0$, 则

$$\nu_1(E) = \nu(E \cap B^c) = \int_{E\cap B^c} f\ d(\mu+\nu) = \int_{E\cap B^c} f\ d\nu.$$

所以

$$\int_{E\cap B^c} (1-f)\ d\nu = 0.$$

由于在B^c上$f < 1$, 所以

$$\nu_1(E) = \nu(E \cap B^c) = 0.$$

于是$\nu_1 \le \mu$. 分解的存在性得证.

下面证明唯一性. 设还有一个满足相同性质的分解

$$\nu = \tilde{\nu}_1 + \tilde{\nu}_2,$$

相应的X的分割为

$$X = \tilde{B} \cup \tilde{B}^c.$$

若$E \in \mathscr{F}$且$E \subset B \cup \tilde{B}$, 有$\mu(E) = 0$, 因而由$\nu_1,\ \tilde{\nu}_1 \le \mu$有

$$\nu_1(E) = \tilde{\nu}_1(E) = 0,$$

因此

$$\nu_2(E) = \tilde{\nu}_2(E).$$

而对于$E \in B^c \cap \tilde{B}^c$, 有

$$\nu_2(E) = \tilde{\nu}_2(E) = 0.$$

于是$\forall E \in \mathscr{F}$, 有

$$\begin{aligned}
\nu_2(E) &= \nu_2(E \cap (B \cup \tilde{B})) + \nu_2(E \cap (B^c \cap \tilde{B}^c)) \\
&= \tilde{\nu}_2(E \cap (B \cup \tilde{B})) + \tilde{\nu}_2(E \cap (B^c \cap \tilde{B}^c)) \\
&= \tilde{\nu}_2(E),
\end{aligned}$$

由此又有

$$\nu_1(E) = \tilde{\nu}_2(E).$$

分解的唯一性得证. 完毕.

推论15.7.3. *存在$f \in L^1(X, \mathscr{F}, \mu)$及$\mu$-可略集$N$使$\forall E \in \mathscr{F}$有*

$$\nu(E) = \int_E f(x)\mu(dx) + \nu(E \cap N).$$

并且上述分解在下面意义下唯一的: 若另有$g \in L^1(X, \mathscr{F}, \mu)$及$\mu$-可略集$M$使$\forall E \in \mathscr{F}$有

$$\nu(E) = \int_E f(x)\mu(dx) + \nu(E \cap M),$$

则$f = g$ *a.e.* 且$(\mu + \nu)(M \triangle N) = 0$.

上述的分解也叫Lebesgue分解.

有了Lebesgue分解之后, 我们就可以推广Radon-Nikodym导数的概念到一般未必绝对连续的测度.

定义15.7.4. *设μ、ν是两个有限测度, ν关于μ的Lebesgue分解为*

$$\nu(E) = \nu(E \cap B) + \int_E f(x)\mu(dx), \quad \mu(B) = 0.$$

定义

$$\frac{d\nu}{d\mu} = \begin{cases} f(x) \text{ 如果} x \in B \\ \infty \text{ 如果} x \in B^c \end{cases}$$

$\frac{d\nu}{d\mu}$称为ν对μ的导数.

由推论15.7.3, $\frac{d\nu}{d\mu}$是$\mu + \nu$-唯一确定的.

我们有

命题15.7.5. *(i)*

$$\frac{d\mu}{d\nu} = \left(\frac{d\nu}{d\mu}\right)^{-1}.$$

(ii)若$\mu_1 + \mu_2 \ll \mu$, 则

$$\frac{d\mu_1}{d\mu_2} = \frac{d\mu_1}{d\mu}\frac{d\mu}{d\mu_2}.$$

现在讨论导数Radon-Nikodym导数的收敛性问题. 设μ与ν均是$\mathscr{F}$上的有限测度, $\mathscr{F}_n$, $n = 1, 2, \cdots$,为$\mathscr{F}$的子σ代数且$\mathscr{F} = \sigma(\cup_{n=1}^{\infty}\mathscr{F}_n)$. μ、ν在$\mathscr{F}_n$上的限制分别记为μ_n、ν_n. 显然, 若$\nu \leq \mu$, 则$\forall n$, $\nu_n \leq \mu_n$——此时我们称ν关于μ局部绝对连续, 记为$\nu \leq^{\text{loc}} \mu$. 这个问题的反问题是: 若$\nu \leq^{\text{loc}} \mu$, 是否无条件地有$\nu \leq \mu$? 若否, 在什么条件下有此结果?

定理15.7.6. *设$f_n = \frac{d\nu}{d\mu}|_{\mathscr{F}_n}$, $f = \frac{d\nu}{d\mu}$, 则*

(i)$\lim_{n\to\infty} f_n = f$, $\mu + \nu$-a.e.;

(ii) $\nu \ll \mu$的充要条件是: $\forall n$, $\nu|_{\mathscr{F}_n} \ll \mu|_{\mathscr{F}_n}$且$\nu(\lim_{n\to\infty} f_n < \infty) = 1$;

(iii)$\nu \perp \mu$的充要条件是$\nu(\lim_{n\to\infty} f_n = \infty) = 1$, 或者等价地$\mu(\lim_{n\to\infty} f_n < 0) = 1$.

由此我们可以得到关于乘积测度的Kakutani定理.

定理15.7.7. *(Kakutani)*设$(X_n, \mathscr{F}_n, \mu_n)$是概率空间, ν_n是$(X_n, \mathscr{F}_n)$上的概率测度且$\nu_n \ll \mu_n$. 令$\mu = \prod_{n=1}^{\infty} \mu_n$, $\nu = \prod_{n=1}^{\infty} \nu_n$, 则要么

$$\nu \ll \mu$$

要么

$$\nu \perp \mu.$$

更精确一些, 记

$$f_n := \frac{d\nu_n}{d\mu_n},$$

则当

$$\prod_{n=1}^{\infty} \int_{X_n} f_n^{\frac{1}{2}} \ d\mu_k > 0 \text{ 或 } = 0$$

时, 分别有$\nu \ll \mu$或$\nu \perp \mu$.

参考文献

[1] Parthasarathy, K. P.: *Probability measure on metric spaces*, Academic press, New York and London, 1967.